

PROFICIENCY AND QUALITY IN FOREIGN LANGUAGE READING:
A STUDY OF THE RELATIONSHIP BETWEEN PROFICIENCY LEVEL
AND READING OUTCOME

Francisco Galicia-Ortega

Ph.D

University of Edinburgh

1988



ABSTRACT

The main purpose of the study reported in this thesis was to explore the relationship between the degree of proficiency in a foreign language and the quality of reading outcomes.

Ninety Mexican university students were the subjects of the study. They were divided into three groups according to their level of command of their foreign language, English: "Beginners," "Intermediate" and "Advanced." They were asked to summarize and to make comments on a text in English. Their performance in the foreign language was compared to the way in which they performed the same tasks in their mother tongue (Spanish). As a control, half the number of subjects took the test in English first and then the test in Spanish. The other half received the tests in the reverse order.

The quality of the subjects' reading outcomes was assessed by two independent judges, using as a basis for the analysis the SOLO Taxonomy ("The Structure of the Observed Learning Outcome") developed by Biggs and Collis (1982). The interjudge reliability coefficient was significant beyond the 0.1% level.

An ANOVA test was applied to the data to determine the statistical significance of the differences observed amongst the contrasting groups. A three-way ANOVA was necessary to analyse the effects of the three independent variables of the study ("Language of Test," "Level of Proficiency" and "Order of Administration") on the dependent one ("Reading Outcome").

The results of the qualitative and quantitative analyses of the data are discussed and conclusions drawn, the main one being that language of test interacts significantly with proficiency level.

DECLARATION

The work presented in this thesis is my own.

Francisco Galicia-Ortega.

CONTENTS

Page

List of Appendices.

List of Tables.

List of Figures.

Acknowledgements.

Introduction. 1

CHAPTER 1. Background. 5

1.1 Introduction. 6

1.2 Background Information about the Language
Situation and Educational System in Mexico. 7

1.3 Reading in a Foreign Language. 12

1.3.1 ✓ Definition of Terms. 12

1.3.2 Research in Reading in a Foreign Language. 13

1.3.2.1 The Roles of Background Knowledge and the
Text; plus some Methodological Aspects. 14

1.3.2.2 The Roles of Linguistic Competence and
Reading Skill. 20

1.4 Evaluating the Quality of Reading Outcomes. 27

1.4.1 A Definition of Quality. 27

1.4.2 Evaluating Quality in Education. 30

1.4.3 Evaluating Reading Outcomes. 33

1.4.3.1 The Brunel Model. 33

1.4.3.2 The Gothenburg Model. 35

1.4.3.3 The SOLO Model. 37

	<u>Page</u>
CHAPTER 2. <u>The SOLO Taxonomy and Reading Outcomes.</u>	41
2.1 Introduction.	42
2.2 Description of the SOLO Taxonomy.	44
2.3 An Application of the SOLO Taxonomy to	
✓ the Analysis of Reading Outcomes: An	
✓ Exploratory Study in L1 and FL.	51
2.3.1 Background.	51
2.3.2 Analysis of the Materials from a SOLO	
Perspective.	52
2.3.3 Application of the Tests.	57
2.3.3.1 Subjects.	57
2.3.3.2 Analysis of Subjects' Answers.	58
2.3.4 Comments.	61
CHAPTER 3. <u>The Experiment.</u>	64
3.1 Introduction.	65
3.2 Research Design.	66
3.2.1 Variables	67
3.2.2 Type of Design.	69
3.2.3 Hypotheses.	71
3.3 Materials.	73
3.3.1 Foreign Language Proficiency Test.	73
3.3.2 Reading Comprehension Tests.	74
3.3.2.1 Text Selection.	75
3.3.2.2 Readability Study.	77

	<u>Page</u>
3.3.2.3 Analysis of Texts 1 and 2 from a SOLO Perspective.	83
3.3.2.4 Specialists' Summaries of Texts 1 and 2.	89
3.3.2.5 Pilot Study.	91
3.4 Experimental Subjects.	97
3.5 Procedure.	101
3.5.1 Application of Tests.	102
3.5.2 Scoring of Tests.	104
CHAPTER 4. <u>Results.</u>	106
4.1 Introduction.	107
4.2 Proficiency Levels.	108
4.3 SOLO Levels.	116
4.3.1 SOLO Scores.	116
4.3.2 Interjudge Reliability.	124
4.4 ANOVA	129
4.4.1 Complete ANOVA.	132
4.4.2 Breakdown Analyses of Variance.	143
4.4.2.1 Breakdown Analysis of Variance by "Language" (of Test).	150
4.4.2.2 Breakdown Analysis of Variance by "Level" (of Proficiency).	157
CHAPTER 5. <u>Discussion of Results.</u>	167
5.1 Introduction.	168
5.2 ANOVA Results and SOLO Levels.	168

	<u>Page</u>
5.2.1 Discussion of the Interaction "Language x Level" from a SOLO Perspective.	169
5.2.1.1 SOLO Levels.	171
5.2.1.2 Categories Resulting from the "Language of Test x Level of Proficiency" Interaction.	190
5.2.2 Discussion of the "Order of Administration" Effect from a SOLO Perspective.	200
5.2.2.1 Overall Variation in SOLO Terms.	200
5.2.2.2 Individual Variation in SOLO Terms.	205
5.3 Main Findings in Relation to Experimental Hypotheses.	211
5.3.1 Order of Administration.	214
5.3.2 "Language of Test x Level of Proficiency" Interaction.	215
CHAPTER 6. <u>Conclusions.</u>	219
6.1 Introduction.	220
6.2 Reliability.	220
6.3 Validity.	221
6.3.1 Adequacy of Task Selection.	222
6.3.2 Adequacy of Text Selection.	223
6.3.3 Adequacy of Research Tools in Relation to Purposes of the Study.	225
6.3.3.1 Qualitative and Quantitative Analyses.	225
6.3.3.2 Norm-Referenced and Criterion-Referenced Evaluation.	227

	<u>Page</u>
6.3.3.3 Hypothesized and Observed Results.	230
6.4 Summary of Main Findings.	231
6.5 Implications.	236
6.5.1 Practical.	236
6.5.2 Future Research.	239
APPENDICES.	241
BIBLIOGRAPHY.	352

APPENDICES

- APPENDIX A. Test A. (Used in Exploratory Study).
- APPENDIX B. Test B. (Used in Exploratory Study).
- APPENDIX C. Transcripts of Responses Obtained in an
Exploratory Application of Tests A and B.
- APPENDIX D. Test 1 (English).
- APPENDIX E. Test 2 (English).
- APPENDIX F. Test 1 (Spanish).
- APPENDIX G. Test 2 (Spanish).
- APPENDIX H. Transcripts of Specialists' Summaries of
Texts 1 and 2.
- APPENDIX I. Instructions Used in Pilot Study in
Tests 1 and 2.
- APPENDIX J. Samples of Subjects' Responses in Tests
1 and 2 (Pilot Study).
- APPENDIX K. Data on Language Proficiency. (Subjects'
ELBA Scores and Classification).
- APPENDIX L. Data on Reading Comprehension. (Subjects'
Scores in Tests 1 and 2).
- APPENDIX M. Tables Used in the Preparation of Main Table
of Analysis of Variance (Complete ANOVA).
- APPENDIX N. Tables Used in the Preparation of the Tables
of Analysis of Variance by "Language" (of Test),
and by "Level" (of Proficiency). (Breakdown
Analyses of Variance).
- APPENDIX O. Transcripts of Subjects' Responses Representative
of Each Level of the SOLO Taxonomy.

TABLES

- Table 1.1 Number of Students in the English Department at the Foreign Language Centre, National University, Mexico, during the Second Semester of the 1986-1987 Academic Year.
- Table 2.1 Basic Features of the SOLO Taxonomy (Based on Biggs and Collis, pp.24-5, Table 2.1).
- Table 2.2 Examples of Responses that May Be Expected in Tests A and B: Their Main Characteristics from a SOLO Perspective.
- Table 3.1 Flesch Reading Ease Scores of Texts 1 and 2, in English and Spanish.
- Table 3.2 Results of the Wilcoxon Test Applied to Different Pairs of the Texts in the Study.
- Table 3.3 NSYLL of Texts 1 and 2, in English and Spanish.
- Table 3.4 W/S of Texts 1 and 2, in English and Spanish.
- Table 3.5 RE of Texts 1 and 2, in English and Spanish.
- Table 3.6 Results of the χ^2 Test Applied to the NSYLL, W/S and RE Values of the Texts in the Study.
- Table 3.7 Main Characteristics of the Responses that May Be Expected in Test 1 (English and Spanish), from a SOLO Perspective.
- Table 3.8 Responses that May Be Expected in Test 2 (English and Spanish), from a SOLO Perspective.
- Table 3.9 Specialists' Scores in Summaries of Texts 1 and 2.

Table 3.10	Tests Taken by Subjects in Pilot Study.
Table 3.11	Marks Allotted to SOLO Levels.
Table 3.12	Subjects' Marks in Test 1: Differences Between Two Judges. (Pilot Study).
Table 3.13	Subjects' Marks in Test 2: Differences Between Two Judges. (Pilot Study).
Table 3.14	Degree of Agreement Between Two Judges on Tests 1 and 2 (Pilot Study).
Table 3.15	Correlation Between Two Judges on Test 1, Test 2, and Tests 1 and 2. (Pilot Study).
Table 3.16	Number of Subjects Participating in the Experiment, by Level of Proficiency (CELE Classification), and Order of Administration of Tests.
Table 3.17	Number of Subjects Who Completed the Battery of Tests in the Experiment, by Level of Proficiency (ELBA Classification), and Order of Administration of the Tests.
Table 3.18	Experimental Subjects, by Level of Proficiency and Order of Administration of Tests.
Table 4.1	ELBA Scores, According to Subjects' CELE Classification.
Table 4.2	Number and Percentage of Subjects in the "Advanced," "Intermediate" and "Beginners" Groups, According to their ELBA Scores.
Table 4.3	Number of Subjects, by CELE Classification and ELBA Scores.

Table 4.4	Data for One-Way ANOVA, ELBA Scores.
Table 4.5	ANOVA for Subjects' ELBA Scores.
Table 4.6	SOLO Scores: Number of Subjects by "Language of Test," "Order of Administration" and "Level of Proficiency."
Table 4.7	Differences Between Initial Scores Awarded by Both Judges in Both Tests.
Table 4.8	Number of Cases of Initial Disagreement of More than One Level: Modified and Unmodified.
Table 4.9	Degree of Agreement Between Two Judges on Test 1.
Table 4.10	Degree of Agreement Between Two Judges on Test 2.
Table 4.11	Degree of Agreement Between Two Judges on Tests 1 and 2.
Table 4.12	Correlation Between Two Judges on Test 1, Test 2, and Tests 1 and 2.
Table 4.13	Table of Analysis of Variance (Complete ANOVA).
Table 4.14	Table of Analysis of Variance. Breakdown ANOVA by "Language:" "English".
Table 4.15	Table of Analysis of Variance. Breakdown ANOVA by "Language:" "Spanish".
Table 4.16	Table of Analysis of Variance. Breakdown ANOVA by "Level" (of Proficiency): "Advanced".
Table 4.17	Table of Analysis of Variance. Breakdown ANOVA by "Level" (of Proficiency): "Intermediate".
Table 4.18	Table of Analysis of Variance. Breakdown ANOVA by "Level" (of Proficiency): "Beginners".

- Table 5.1 SOLO Scores Taking Interaction "Language x Level" into Account: Number (and Proportion) of Subjects per Category.
- Table 5.2 SOLO Scores as a Function of "Order of Administration:" Number (and Proportion) of Subjects per Category.
- Table 5.3 Individual Variation in SOLO Scores from Test 1 to Test 2: Number (and Proportion) of Subjects per Category.
- Table 5.4 Relationship Between Experimental Hypotheses, Independent Variables and Main Factors in Three-Way ANOVA Test.

FIGURES

- Figure 2.1 Answers to the Question "Why is the Side of the Mountain that Faces the Coast Usually Wetter than the Side Facing the Interior?" (Used by the Authors as an Example of the Way in Which the SOLO Taxonomy Can be Used in Qualitative Analyses).
(Biggs and Collis, pp. 4-5).
- Figure 2.2 The Complete SOLO Taxonomy, as Applied to History Items by Its Authors (Biggs and Collis, p.36).
- Figure 3.1 Schematic Representation of Research Design.
- Figure 3.2 Flesch Readability Formula.
- Figure 4.1 SOLO Scores: Beginners (n=15), Test 1 (English).
- Figure 4.2 SOLO Scores: Intermediate (n=15), Test 1 (English).
- Figure 4.3 SOLO Scores: Advanced (n=15), Test 1 (English).
- Figure 4.4 SOLO Scores: Beginners (n=15), Test 2 (English).
- Figure 4.5 SOLO Scores: Intermediate (n=15), Test 2 (English).
- Figure 4.6 SOLO Scores: Advanced (n=15), Test 2 (English).
- Figure 4.7 SOLO Scores: Beginners (n=15), Test 1 (Spanish).
- Figure 4.8 SOLO Scores: Intermediate (n=15), Test 1 (Spanish).
- Figure 4.9 SOLO Scores: Advanced (n=15), Test 1 (Spanish).
- Figure 4.10 SOLO Scores: Beginners (n=15), Test 2 (Spanish).
- Figure 4.11 SOLO Scores: Intermediate (n=15), Test 2 (Spanish).
- Figure 4.12 SOLO Scores: Advanced (n=15), Test 2 (Spanish).
- Figure 4.13 Analysis of Variance: Kerlinger's Equation Used in the Study (Kerlinger, 1967, p.231).

- Figure 4.14 Equation to Obtain the Between-Subjects Sum of Squares.
- Figure 4.15 Graphic Representation of the "Order" (of Administration) Main Effect: English Tests.
- Figure 4.16 Graphic Representation of the "Order" (of Administration) Main Effect: Spanish Tests.
- Figure 4.17 Graphic Representation of the Interaction "Language x Level." (Results of the First Administration of the Tests).
- Figure 4.18 Graphic Representation of the Interaction "Language x Level." (Results of the Second Administration of the Tests).
- Figure 4.19 Graphic Representation of Lack of Interaction Between the Variables "Language" (of Test) and "Level" (of Proficiency): First Administration. (Hypothetical).
- Figure 4.20 Graphic Representation of Lack of Interaction Between the Variables "Language" (of Test) and "Level" (of Proficiency): Second Administration. (Hypothetical)
- Figure 4.21 Reading Outcome as Determined by the "Beginners," "Intermediate" and "Advanced" Levels of the "Level" (of Proficiency) Variable, and by the "English First" and "English Second" Levels of the "English" Variable.
- Figure 4.22 Reading Outcome as Determined by the "Beginners," "Intermediate" and "Advanced" Levels of the "Level" (of Proficiency) Variable, and by the "Spanish First" and "Spanish Second" Levels of the "Spanish" Variable.

- Figure 4.23 Reading Outcome as a Function of "Language" (of Test) and "Order" (of Administration), for the "Advanced" Group.
- Figure 4.24 Reading Outcome as a Function of "Language" (of Test) and "Order" (of Administration), for the "Intermediate" Group.
- Figure 4.25 Reading Outcome as a Function of "Language" (of Test) and "Order" (of Administration), for the "Beginners" Group.
- Figure 5.1 Prestructural Responses as a Function of "Language x Level."
- Figure 5.2 Transitional Responses (from Prestructural to Unistuctural) as a Function of "Language x Level."
- Figure 5.3 Unistuctural Responses as a Function of "Language x Level."
- Figure 5.4 Transitional Responses (from Unistuctural to Multistuctural) as a Function of "Language x Level."
- Figure 5.5 Multistuctural Responses as a Function of "Language x Level."
- Figure 5.6 Transitional Responses (from Multistuctural to Relational) as a Function of "Language x Level."
- Figure 5.7 Relational Responses as a Function of "Language x Level."
- Figure 5.8 Transitional Responses (from Relational to Extended Abstract) as a Function of "Language x Level."
- Figure 5.9 Extended Abstract Responses as a Function of "Language x Level."

- Figure 5.10 SOLO Levels as a Function of the Interaction "Language x Level:" English Test, Beginners Group.
- Figure 5.11 SOLO Levels as a Function of the Interaction "Language x Level:" Spanish Test, Beginners Group.
- Figure 5.12 SOLO Levels as a Function of the Interaction "Language x Level:" English Test, Intermediate Group.
- Figure 5.13 SOLO Levels as a Function of the Interaction "Language x Level:" Spanish Test, Intermediate Group.
- Figure 5.14 SOLO Levels as a Function of the Interaction "Language x Level:" English Test, Advanced Group.
- Figure 5.15 SOLO Levels as a Function of the Interaction "Language x Level:" Spanish Test, Advanced Group.
- Figure 5.16 SOLO Scores as a Function of "Order of Administration." Graphic Representation of Percentages of Subjects per SOLO Level, in First and Second Tests.
- Figure 6.1 Relationship Observed Between the Variables in the Study.

ACKNOWLEDGEMENTS

I would like to express my gratitude to the people and institutions who made this study possible.

To my supervisor, Dr. Alan Davies, for his wise advice and encouragement.

To my wife who, in spite of being engaged in her own research, found the time to act as a second translator, reading specialist and marker.

To my colleagues at the Foreign Language Teaching Centre (CELE) of the National Autonomous University of Mexico (UNAM) who helped and encouraged me throughout the various stages of the study.

To the students of the English Department of the Centre (CELE) who participated in the experiment.

To my friends in Edinburgh who participated in the exploratory and pilot studies.

To The Macmillan Publishers Ltd., for their kind permission to use excerpts from the text "The International Community and the Environment," by M. F. Strong, in the reading comprehension tests.

To the Department of Applied Linguistics, University of Edinburgh, for their kind permission to use selected sections of the English Language Battery Test (ELBA Test).

To the British Council and the National Autonomous University of Mexico for their financial support.

To Mrs E. McCormack for preparing the final typescript.

To my daughter

María Amparo

INTRODUCTION

Introduction

The main purpose of this study is to explore possible relationships between the degree of proficiency in a foreign language and the quality of reading outcomes. After surveying the relevant literature, an experiment designed to test the proficiency-reading outcome relationship is described; and its results, reported and discussed.

In this introduction, after a brief account of the experiment in question an overview of the way in which the thesis has been structured is provided.

Let us consider the experiment first. Ninety Mexican university students with different levels of command of English - their foreign language - were the subjects of the study. They were asked to summarize and to make comments on a text in English. These reading tasks were chosen because they promote receptive and responsive reading styles (Pugh, 1978, pp.,52-5). It was felt that the likely outcomes of these reading styles could provide a suitable source of information for the purpose of this study; namely, to analyse qualitatively the reading outcomes.

The students' responses were, consequently, analysed qualitatively. The basis for this analysis was provided by the SOLO Taxonomy (SOLO being an acronym for "The Structure of the Observed Learning Outcome") developed by Biggs and Collis (1982).

In order to ensure a better understanding of the results, the subjects' performance in the foreign language was compared to the way in which they performed the same tasks in their mother tongue (Spanish).

As a further control, half the number of subjects took the test in English first and then the test in Spanish. The remaining half received the tests in the reverse order.

The thesis is structured as follows: Chapter One establishes the background context of this investigation. The educational context is examined first. Then, regarding the theoretical framework, two aspects are considered: the relevant research on reading comprehension in a foreign language, and the problem of analysing the quality of reading outcomes.

Chapter Two explores in depth the way in which the SOLO Taxonomy can be applied to the analysis of the quality of reading outcomes, both in the mother tongue and in a foreign language. An account is given of an exploratory study and its relation to the main experiment (reported in the following chapters).

Chapters Three to Five deal with the main experiment. Chapter Three includes a discussion of the rationale behind the experiment, and a description of the experimental design, subjects, materials and procedure. More attention is given to the section dealing with the materials used: it includes information about the English proficiency test used, it provides a complete description of the English and Spanish reading comprehension tests, and apart from a discussion of the main problems posed by the design of the tests, this section also discusses the results of a pilot study, and the marking scheme developed on the basis of the SOLO Taxonomy.

Chapter Four reports the results of the experiment. First, the way in which the subjects performed in the foreign language

proficiency test is dealt with. Then, the quality of the subjects' reading outcomes is assessed by two independent judges. After establishing the degree of interjudge reliability, the subjects' scores on the reading tests in English and Spanish are analysed by means of an ANOVA test. The results of this test are included in the last section of this chapter.

The discussion which follows in Chapter Five relates the results of the three tests (Proficiency in English, Reading Comprehension in English, and Reading Comprehension in Spanish) to the hypothesized outcomes of the experiment.

The last chapter of the thesis presents the conclusions reached. A summary of the main findings follows a discussion of the adequacy of the materials used, and on the whole, of the experimental design. In the final section, the implications of the study are considered.

CHAPTER ONEBackground

Chapter One

Background

1.1 Introduction

This chapter attempts to provide the necessary information to establish the context in which the present study originated and was carried out.

Accordingly, the following section presents relevant information regarding the language situation and educational system in Mexico; more specifically about foreign language teaching and research at the National University.

Then, consideration is given as to how the study of reading comprehension in a foreign language for academic purposes at university level has been approached. Different aspects of reading in a foreign language have already been chosen as relevant topics of study. For instance, the role played by the reader's linguistic competence or cultural knowledge, or even problems posed to the reader by a different writing system to his own, have been studied. (Cf., for example, Alderson and Urquhart, 1984; Barnitz, 1985; Bruder and Henderson, 1986; Pugh and Ulijn, 1984; and Ulijn and Pugh, 1985). Of particular relevance to this study is how the linguistic competence aspect has been approached, and will therefore be the main concern of the section in question.

In the final section of the chapter consideration is given to some ways in which the problem of assessing learning qualitatively has been dealt with (cf., for instance, Biggs and Collis, 1982; Bloom, 1956; Entwistle, 1981; Entwistle and Ramsden, 1983; Marton, Hounsell and Entwistle, 1984; and Pollitt, et al., 1985), and

it is argued that Biggs and Collis's SOLO Taxonomy can be used as a suitable tool for analysing the quality of reading outcomes.

1.2 Background Information about the Language Situation and Educational System in Mexico

Stern has pointed out that "shifts in emphasis on French, English, German, Spanish, Portuguese, Russian, or Dutch as second languages throughout the world have mirrored the ups and downs of political and economic power and prestige." (Stern, 1983, p.278). This observation can be taken as a starting point to establish the context in which the present study fits. It seems to me that it is also applicable to the case of foreign language teaching, and that the development of ESP (English for Specific Purposes) is a relevant example of this.

In his survey of English language teaching, Howatt discusses how "the rapid development in the seventies of English for Specific Purposes" was due to the fact that by 1970, English had already played the role of "the lingua franca of modern science and technology" for some years, and "was also the language of trans-national commerce, finance, and practical communication generally." (Howatt, 1984, pp 221-2).

Regarding the foreign language teaching situation in Latin America, Alvarez's analysis supports the foregoing observations. He examines the way in which social, political and economic influences from Europe - and more recently from the United States of America - have influenced foreign language teaching policies in the area. He points out some major similarities among the different countries

in the region. For example, during the first half of this century French was the main foreign language in Latin America, while English was second, sometimes contending with German or Italian. Then, after the Second World War the situation changed, English becoming the main foreign language. This coincided, of course, with the fact that English was gaining preponderance all over the world in science, commerce and technology. Although Alvarez is basically concerned with discussing the teaching of French, the main reason he sees for teaching it in Latin America nowadays, is in agreement - it seems to me - with current views of English teaching as well. That is, that the teaching of a foreign language should serve the purpose of providing Latin American countries with an efficient tool to fight economic, technological, scientific and cultural underdevelopment. (Alvarez, 1980).

As far as the Mexican situation in particular is concerned, it should be noted that Spanish is the mother tongue of the vast majority of the population. It is the official language of the country, and the main language of instruction.

English, on the other hand, is basically taught as a foreign language. With the exception of some schools in the private sector, it is not included in the curriculum at the primary level. At the secondary and preparatory levels (roughly equivalent to Forms S1 through S6 in Scotland), it is the main foreign language taught. At university level, foreign language instruction is generally excluded from the curriculum, though a foreign language examination is part of the degree requirements for most courses of study. English, once again, tends to be the foreign language favoured.

With the advent of ESP in the seventies, the foreign language situation at the tertiary level began to be systematically explored within the main universities of the country, notably the National Autonomous University of Mexico (UNAM), and the Metropolitan Autonomous University (UAM). The general trend identified meant that the use of a foreign language at the university level was convenient rather than essential, despite what the official requirements seemed to indicate. The following considerations are representative of the general pattern that emerged regarding English:

The needs in question have seemed to be greatest in terms of bibliographical research within the university program itself and during lifetime career development. It is not so much whether English is indispensable or not. In the university we already know that with the possible exception of particular majors, such as in the case of Anthropology, students can make their way to a degree without it. Career life, moreover, may be equally impervious to English for the majority. But...academic and career life can receive a real boost in many cases, by the study of English. About half the books in the UAM library alone are in English, although generally these are not books used as texts but are rather consulted for research development. In like manner, specialized texts and journals publishing the latest research findings are frequently found in Mexico only in the English original or the English translation. . . Students and teachers at the UAM have given support to our view that bibliographical research is the prime need for English among Mexican university students and academicians.... (Connaughton and Flores, 1978, pp.23).

Let us now turn our attention to the National Autonomous University of Mexico (UNAM), the largest university of the country (in January 1986, the number of matriculated students, including those at the preparatory level, was 256,693). In fact, the university is divided into several campuses. Each one of these campuses has a Foreign Language Centre.

The number of students in the main campus (The University City Campus) was 87,501, in January 1986. In this campus, besides the main Foreign Language Centre (The CELE, or Centro de Enseñanza de Lenguas Extranjeras), there are foreign language facilities in several of the schools running backup courses for their own students. Such courses are usually language courses for academic purposes. Of these, the majority aim at improving the ability of the students to read specialised materials in English.

At CELE, on the other hand, although there are also some courses devoted to reading for academic purposes (in English, French, German and Italian), most language courses have as objective to develop overall proficiency. At present, the Centre offers courses in twelve languages: Arabic, Bulgarian, Chinese, English, French, German, Hebrew, Italian, Japanese, Portuguese, Russian and Swedish. The population fluctuates around 10,000 students per year, around 5,000 each semester.

Of the languages taught, English is again the one that attracts the larger number of students. In the second semester of the academic year 1986-1987, semester in which the experiment reported here was carried out, there were 2,059 students in the English Department. (cf. Table 1.1).

LEVEL	NUMBER OF STUDENTS	NUMBER OF CLASSES
1st	510	19
2nd	477	16
3rd	305	12
4th	284	10
5th	218	8
6th	107	6
7th	26	1
8th	43	2
9th	21	1
Reading Comp.	68	2
TOTALS	2059	77

Table 1.1 Number of students in the English Department at the Foreign Language Centre, National University, Mexico, during the second semester of the 1986-1987 academic year.

Important as the language courses in themselves may be, the CELE has other no less important functions as well. These include the development of teaching materials and examinations for the different languages taught, the running of a Teacher Training Course (in English, French, German, Italian and Portuguese), and an M.A. programme in Applied Linguistics (with reference to the teaching of English, French or German). The Centre also offers advisory services to other schools within the university, and to other institutions in the country.

Also, as has already been mentioned, some research has been carried out in the Centre since the seventies. The initial research projects were basically concerned with looking for solutions to the more immediate problems within the university (e.g. needs analyses of different groups of students, and the design of suitable backup courses). More recently, however, the scope has been widening, and more theoretical aspects of the language teaching/learning processes have received consideration.

One of the topics that has received more attention is reading comprehension for academic purposes, mostly in English (cf. Alderson, 1978; Alderson and Alvarez, 1977; Alderson and Richards, 1977; Alderson and Williamson, 1978; Alderson, Bastien and Madrazo, 1978; Alderson, Bastien and Richards, 1978; Alvarez and Mackay, 1977; Bastien, 1983; Castaños, 1979; Chasan, 1981; Chasan and Williamson, 1979; Galicia, 1983; Mackay and Klassen, 1975; Mackay, et al., 1975 a; 1975 b; 1975 c; 1976 a; and 1976 b). It is in this educational context that the question raised in this study emerged. Its antecedents regarding research in foreign language reading will be examined in the next section.

1.3 Reading in a Foreign Language

1.3.1 Definition of Terms

Since the main concern in this study is reading in a foreign language, it will be convenient to start this section with the definition of "foreign language" that has been adopted here, which is that one found in the Longman Dictionary of Applied Linguistics: "a language which is not a native language in a country. A foreign language is usually studied either for communication with foreigners who speak the language, or for reading printed materials in the language." (Emphasis added). (Richards, Platt, and Weber, 1985, p 108).

The following clarification is made immediately afterwards: "in North American applied linguistics usage, 'foreign language' and 'second language' are often used to mean the same in this sense," whereas in British usage the distinction below is often made:

- (a) a foreign language is a language which is taught as a school subject but which is not a medium of instruction in schools nor as a language of communication within a country (eg., in government, business or industry). English is described as a foreign language in France, Japan, China, etc.
- (b) a second language is a language which is not a native language in a country but which is widely used as a medium of communication (eg., in education and government) and which is usually used alongside another language or languages. English is described as a second language in countries such as Fiji, Singapore, and Nigeria.

(op. cit., pp 108-9).

As for a definition of reading, it seems to me that the one given by Barnitz is a clear statement of current views of the reading process. (cf., for instance, Alderson and Urquhart, 1984, pp xv-xxviii; Bruder and Henderson, 1986, pp.3-10; Hall, White, and Guthrie,

1986, pp.90-1; Nuttall, 1982, ppl-18; and Ulijn, 1984, pp.66-8), and that it can serve as a basic framework for the purposes of this study. Barnitz states:

Reading is a complex communication process in which the mind of the reader interacts with the text in a particular setting or context. During the reading process, readers construct a meaningful representation of text through an interaction of their conceptual and linguistic knowledge with the cues that are in the text.

(Barnitz, 1985, p.3).

It seems to me that this view of the reading process can be taken regarding reading in the mother tongue as well as reading in a foreign language, for as Alderson and Urquhart put it, "it is not clear to what extent reading in a foreign language is different from reading in a first language." (op cit., p. xv).

1.3.2 Research in Reading in a Foreign Language

The study of reading comprehension of academic texts in a foreign language still offers a rich field of inquiry to the applied linguist. The following statement is representative of the awareness of this fact found in the relevant literature: "... within the vast literature on reading, there is relatively little on reading in foreign languages, less on reading for specific purposes..." (Pugh and Ulijn, 1984, p.iv).

In what there is, however, one could draw a first distinction between two groups of authors. On the one hand, some have concentrated their attention on discussing whether reading in a foreign language is fundamentally a reading or a language problem, to use Coady's (Coady, 1979, p.9), and Alderson's terms (Alderson 1984). On the other hand, one could include in the second group those studies that have focused on other - either

different or more specific - aspects of the nature of reading in a foreign language.

Although of importance in themselves, the studies in the second group are not directly relevant to the present investigation. Therefore, I will mention in the sub-section that follows some of such studies just to illustrate other types of problem that have been of interest to authors working in the area.

1.3.2.1 The Roles of Background Knowledge, and the Text; plus some Methodological Aspects

Some variables affecting reading comprehension in a foreign language that have already been submitted to exploratory experimentation have to do with the reader's background knowledge. So, for instance, in the abstract of a study carried out in this area, Carrell states:

Research in native (English) and nonnative (ESL) reading comprehension has shown that the ability to understand texts is based not only on the comprehender's linguistic knowledge, but also on general knowledge of the world and the extent to which that knowledge is activated during processing.
(Carrell, 1983, p.183).

Different types of background knowledge have been studied. One has been cultural knowledge, the knowledge of "the values and patterns of behaviour" of a given social group. (Cf. Steffensen and Joag-Dev, 1984, P.53). These authors report a study in which subjects from India and the United States were asked to read and recall two texts describing an Indian and an American wedding. Another example of this type of research is that of Johnson (Johnson, 1982), who asked university students of twenty-three nationalities

to read a text in English about Halloween. In both studies, relevant cultural knowledge was found to aid comprehension of the text.

Another type of background knowledge that has received attention is "conceptual knowledge." Alderson points out that for Ulijn and Kempen (Ulijn and Kempen, 1976, cited in Alderson, 1984) conceptual knowledge had to do with "both, the reader's 'knowledge of the text's subject area' ... and knowledge of word meanings, particularly content words rather than function words." (Alderson, 1984, p.13). This view of conceptual knowledge is found again in Ulijn (Ulijn, 1984). Another example of studies in this area is that of Perkins and Jones, who were interested in the effect of preknowledge of the topic of a passage on the comprehension process when reading in English as a Second Language. (Perkins and Jones, 1985).

✓ Cultural and conceptual knowledge are variables that have to do with the reader, and as such have only relatively recently begun to be explored. On the other hand, there is another group of variables - text-bound variables - the study of which can be said to be more deeply rooted in the linguistic tradition, but which nonetheless have a place of their own in current research on reading in a foreign language. The studies that follow can be cited as examples of this type of research.

Ulijn points out that research in this area has focused mainly on lexical and textual levels (Ulijn, 1985, P.23). A study by Alber-DeWolf is an instance of interest on the lexical level. In her conclusion she states that "the ability to scan or search articles in foreign languages should be the first aim in FL reading

courses for scientists," and that this ability "presupposes an understanding of the foreign terminology." Accordingly, "a knowledge of the most productive morphological and morphosyntactic term-formation processes" can be of great help "to decode the foreign terminology." (Alber-DeWolf, 1984, p.173).

Interest on the syntactic level can be illustrated by a study by Berman (Berman, 1984), whose main interest - as pointed out by Alderson and Urquhart - "is in factors affecting readers' ability to parse sentences into their main constituents, and thus derive meaning from these sentences." (Alderson and Urquhart, 1984, P.157).

✓ At the rhetorical level, and as Urquhart points out, "attempts to measure the ways in which the organization of a text affects readers' interpretation of it" became "an academic growth industry in the 1970's." (Urquhart, 1978, p.25). Examples of this line of research in foreign language reading follow: Urquhart has explored the effect of two organizational principles, time and space order, on readability both in LI and in FL, and found similarities between them. (Urquhart, 1984). Carrell, on the other hand, designed a study to find out whether ESL reading can be facilitated by explicit teaching of text structure, and concluded that it can. (Carrell, 1985).

Another group of text-bound variables that has been considered ✓ has to do with decoding problems. Bruder and Henderson, for example, have examined various problems faced by literate adults, who are non literate in English, when they start reading in this language. (Bruder and Henderson, 1986). Regarding the effect of punctuation on beginning reading in English, they refer -

among others - to a study carried out by Nash. Nash studied the effect of punctuation in English on the reading of Arabic and Spanish speaking learners of English. (Nash, 1983, cited in Bruder and Henderson, 1986). A related study is that by Leong, who examines differences between reading an alphabetical script (English) and a morphemic one (Chinese). (Leong, 1984).

The following study will serve to illustrate a different area of inquiry within this group of studies dealing with text-bound variables. Grundin, et al., explored the effect of using translated vs. non-translated texts on the reading of English, Swedish and Japanese speaking children. They found that their subjects did better when reading non-translated texts. (Grundin, et al., 1981).

Within the wide range of studies that I have arbitrarily grouped together in this sub-section, I will include - besides those dealing with the reader's background knowledge, and those concerned with more text-bound variables - studies illustrative of some important methodological aspects of research in the area.

Samuels and Kamil point out that "what has come to be known as the psycholinguistic perspective ... pushed the field to consider underlying assumptions about basic processes in reading." (Samuels and Kamil, 1984, p.185). This perspective was developed within the field of reading research in LI. Its underlying model of the reading process is that of Goodman, which conceives of reading as "a psycholinguistic guessing game" (Cf., for example, Goodman, 1973a; 1973b; Smith, 1973; and 1978). This model has been very influential. Samuels and Kamil, for example, remark in their study of models of the reading process, that it has had - among all earlier and later models - "the greatest impact on conceptions

about reading instruction, particularly early instruction. So strong has been this impact that it is not uncommon to hear or read about THE psycholinguistic approach to reading..." (Samuels and Kamil, 1984, p.187).

This line of research has also influenced foreign language reading research. Its main research tool, the analysis of oral reading miscues, has been used in several studies of foreign language reading. (Cf., for instance, Hudelson, 1981). The following are some examples.

Rigg analysed the reading miscues of subjects from different native languages while reading in English, finding similarities between them. (Rigg, 1977). In a study carried out by Clarke, the Spanish and English reading performances of a good LI reader and a poor LI reader were also analysed using oral miscue procedures. (Clarke, 1979). Cziko, again using miscue analysis, studied the reading strategies of subjects reading in their mother tongue (English), and in a foreign language (French). (Cziko, 1980). Finally, Devine compared miscues of adults learning to read in English as a foreign language to those observed in children learning to read in their mother tongue (English, too); she found similarities between the two groups of subjects. (Devine, 1981).

✓ A new development in second language research which has already influenced foreign language reading research is the use of introspective/retrospective procedures of data collection.

This approach has already been extensively used in LI reading research in, for example, Brunel University (cf., e.g., Augstein, 1976; Augstein and Thomas 1976, and 1984; Augstein, Smith, and Thomas, 1982; Thomas and Augstein, 1976a, and 1976b), and in Gothenburg University (cf., e.g., Fransson, 1977, and 1984; Marton,

Hounsell, and Entwistle, 1984; Marton and Säljö, 1976a, and 1976b; and Svensson, 1977).

Within the field of foreign language reading research, the following works can be cited. Cohen (Cohen, 1984) argues in favour of using "mentalistic measures" to assess how materials are used in silent reading. By means of such measures, the reader can become an informant about - among other things - the strategies used and the difficulties found while reading - Hosenfeld has used introspective/retrospective techniques to discover strategies used by her subjects to solve problems of understanding foreign language texts. (Cf. Hosenfeld, 1984). Markham, also using interviewing techniques developed by Cohen and Hosenfeld (Cohen and Hosenfeld, 1981) explored cloze sensitivity to global comprehension with native English speaking students and ESL students as subjects. (Markham, 1987).

Besides the cloze test, other research instruments that have been used alongside introspective/retrospective techniques are "prediction protocols" and "pause protocols". Henzell-Thomas used "prediction protocols" to explore the interaction of top-down and bottom-up processes in foreign language reading. (Cf., Henzell-Thomas, 1985). While in this study the subjects were asked to read chunks of text and then to predict what was bound to come next, in the "pause protocols" study by Cavalcanti, the subjects were required to read silently "and to think aloud whenever they noticed a pause in their reading process The readers were, therefore, asked to monitor their silent reading process and to start reasoning aloud on the basis of the pauses detected." (Cavalcanti, 1987, p.238).

The last study I would like to include in this group was carried out by Haastrup with the aim of identifying procedures in the decoding of lexical items unknown to the subjects. In it, "pair thinking aloud" was used. This technique involved a pair of learners of English reading an English text in which an unknown word appeared. The subjects had to make guesses about its meaning and agree on a suitable translation to their mother tongue. Their discussion was recorded and analysed. (Haastrup, 1987).

In the next sub-section, I would like to go back to those studies that have been concerned with the discussion of the roles played by linguistic competence and reading skill in foreign language reading.

1.3.2.2 The Roles of Linguistic Competence and Reading Skill

Regarding the roles played by language proficiency and reading ability in foreign language reading, three main positions can be identified in the literature. (Cf., for example, Alderson, 1984; Barnitz, 1985; and Statman, 1987). On the one hand, reading ability has been seen as largely a function of language proficiency. In a second position, the learner's reading problems are basically seen as a reading problem. A third, and more realistic view - it seems to me - conceives the problem as resulting from the interaction of linguistic and reading factors.

These three positions will be briefly discussed now, making reference to studies representative of each.

According to the first position, then, the learner's reading problems in the foreign language are basically explained in terms of his deficient knowledge of the foreign language.

The initial problems posed by reading in a foreign language with a different type of script of one's own - though important in themselves (cf. Bruder and Henderson, 1986) - will not be dealt with here, since they are not problems faced by the readers under consideration in this study.

Two types of deficiency in the learner's knowledge of the foreign language that have been singled out as highly problematic are the lexical and the syntactical ones. Yorio, for example, sees reading difficulties as basically stemming from the learner's inability to handle vocabulary in the foreign language text (Yorio, 1971). On the other hand, syntactic problems become of paramount importance for Cowan, who argues that the reader's knowledge of his native language structures will determine to a great extent the way in which he will deal with the foreign language. The structural differences of the latter can thus be problematic. The lower the reader's proficiency in the foreign language, the more difficulties he may have due to the inadequacy of relying basically on his native language structures. (Cowan, 1976).

Let us now briefly consider the influence that discourse analysis had on the way the foreign language reader's problems were viewed.

Lautamatti (1978) noticed that "answers to problems in the teaching of reading (were being) sought in a more varied knowledge of the foreign language, and not in an understanding of the kind of activities reading comprehension involves". (p.95). She points out, for example, that

with an increasing knowledge of the properties of texts, the teaching of reading has more and more acquired features of discourse analysis: it includes examination of intersentential features such as reference coherence, cohesion, and use of connectives. The influence of socio-linguistics is seen in the application of the theory of speech acts to the study of written discourse,...

(Ibid.)

The following statement can be considered representative of the approach in question:

the difficulties which students encounter arise not so much from a defective knowledge of the system of English but from an unfamiliarity with English use and consequently their needs (must) be met by a course which ... develops a knowledge of how sentences are used in performance of different communicative acts. (Allen and Widdowson, 1974, cited in Coulthard, 1977, p.148).

Lautamatti, questioning the usefulness of the discourse analysis approach to the teaching of reading in a foreign language, states: "practice emphasizing discourse features may lead to mechanical analysis with no increase in the comprehension of the contents." (op.cit., p.99). This criticism was also shared by Coulthard (op.cit., pp.138-153).

These criticisms can be seen as falling in line with the second position identified in this section regarding the roles played by language proficiency and reading ability in foreign language reading, namely, the one that sees the reader's problems as largely a reading problem.

Such a position finds unqualified expression in Coady's words who, having observed that "alarming numbers of students might have a great deal of proficiency in English and yet read very slowly and with poor comprehension," asserts "we have a reading problem and not a language problem." (Coady, 1979, p.9). An extreme consequence of his view on the independence of language ability and reading is stated - in not very fortunate terms, though - when he remarks that

Conceptual abilities are important in reading acquisition. We can notice this especially in adult foreign students who may fail to achieve the competence necessary for university instruction because they lack intellectual capacity and not totally or necessarily because they cannot learn English.

(op.cit., p.7)

More recently, support to the view that the role of reading ability can override linguistic deficiencies is found in, for example, Beaugrande (1984), and Block (1986). Beaugrande argues in favour of a model of foreign language reading in which native and foreign language reading are taken as more closely related from a processing point of view. Similarities between the two languages can be used to enhance processing. In this paper, he gives examples of techniques that can be used in this respect for English speaking students of German. Block reports a study comparing the comprehension strategies used by ESL college students of different native languages with those of native speakers of English, both classified as nonproficient readers. Among her conclusions she states that language background did not seem to account for the results she found: "the development of strategy use, particularly as it is applied to informative text, does not seem to depend on language-specific features." (Block, 1986, p.485).

Let us now turn our attention to a third possibility of approaching the problem. Here, the importance of the linguistic and reading variables is acknowledged, but the way in which they may affect each other becomes the issue.

A well-known exponent of this position is Clarke who, after studying whether proficient readers in LI transfer their skill to the target language, states:

There is some transfer of skills, for the good readers perform better than the poor readers in both languages, but limited language proficiency appears to exert a powerful effect on the behaviours utilized by the readers. The results of these studies suggest that, while some form of the 'universal hypothesis' may be justified, the role of language proficiency may be greater than has previously been assumed: apparently, limited control over the language 'short circuits' the good reader's system, causing him/her to revert to poor reader strategies when confronted with a difficult or confusing task in the second language.
(Clarke, 1980, p.206)

Another study carried out from this perspective at the time, was that of Hauptman, exploring the roles of syntactic vs. semantic cues and the similarities and/or differences between LI and L2 reading strategies (Hauptman, 1981).

At the National University of Mexico, an interest in this line of research can be found in Alderson, Bastien and Madrazo (1977), whose study leads them to suggest that

the student's knowledge of English is much more important to his comprehension of English texts than is his reading ability in Spanish. Although the finding is not surprising, since one would not expect students with poor English to be able to read well in English, it is important because it appears to provide some sort of answer to those who feel that reading ability in one's mother tongue is the most important determinant of reading ability in English.

(Alderson, Bastien and Madrazo, 1977, p.6)

More recently, Alderson has stated:

To return to the question... is foreign language reading a language problem or a reading problem? The answer, perhaps inevitably, is equivocal and tentative - it appears to be both a language problem and a reading problem, but with firmer evidence that it is a language problem, for low levels of foreign language competence, than a reading problem. As has been suggested, we do not know this yet, and the question needs further refinement and intense investigation.

(Alderson, 1984, p.24)

Other recent examples of the interest to explore possible ways of interaction between language proficiency and reading ability can be found in Barnitz (1985), Laufer and Sim (1985), and Kozminsky and Graetz (1986). So, the experiment reported by Laufer and Sim was carried out to determine the kind of reading that can be expected when the student's knowledge of the foreign language is very incomplete, while Kozminsky and Graetz examined the reading of advanced students of English. On the other hand, Barnitz's monograph includes a more comprehensive discussion of the extent to which language proficiency or language differences may affect reading and learning to read in a foreign language, as well as

possible teaching implications.

As mentioned above, Alderson has pointed out that it is necessary to carry out "intense investigation" in this area. This need has been felt by other people working in the field. For example, at the National University of Mexico Alvarez, Chasan and Galicia (1980), and Bastien (1983), felt this need, basically due to its pedagogical implications. Ulijn too states that it is necessary to examine more closely from this perspective the way in which native language reading may be transferred to foreign language reading. (Ulijn, 1984, pp.76-7). More specifically, Berman points out the convenience of exploring what happens at different levels of proficiency (Berman, 1984, p.152); while Cohen points out the need to study what happens when the same text is read in both languages. (Cohen, 1984, p.184). Statman goes further in this respect, and suggests that the same individual carry out the same task. (Statman, 1987, p.299).

Before closing this sub-section, it will be pointed out that it is possible to find reports of studies exploring the interaction of linguistic proficiency and reading ability with designs that provide different perspectives. For example, a simple procedure that has sometimes been used is to compare two groups of subjects: one of native speakers of the foreign language, and another of non-native ones reading in that language. (Cf., Connor, 1984; Block, 1986; Markham, 1987). Some other times, the same subjects' reading in their mother tongue and in the foreign language has been the object of the study. (Cf., Alderson, Bastien and Madrazo, 1977; Clarke, 1979; and 1980; Kozminsky and Graetz, 1986; and Statman, 1987).

On the other hand, sometimes the subjects' difference of proficiency in the foreign language has been considered as a variable in

the design. Ulijn, however, in a recent "state of the art" survey, found that out of forty-six papers he examined, the level of previous knowledge of the foreign language had been mentioned in only eight; and that of these, only two had included beginning, intermediate and advanced students (Ulijn, 1985, p.19). In the studies reported by Cooper (1984), and Laufer and Sim (1985), the subjects were of two and three different levels of proficiency in the foreign language, respectively.

It is more difficult to find reports of studies bringing together both aspects of the problem, i.e., that compare subjects of different levels of proficiency in the foreign language, reading in both, the native and the foreign languages. (Cf., however, Hauptman, 1981; and, Tudor, 1986).

It is within this framework that the present study falls. As has already been stated, its main objective is to compare the reading performance of subjects with varying levels of foreign language competence to their performance as readers in their mother tongue. By so doing, it is an attempt to respond to the "need for further research performance of foreign language competence and first-language reading ability, on particular tasks..." (Alderson, 1984, p.24).

Besides giving the antecedents to my study regarding foreign language reading research, I will also take into consideration - in the next section of the chapter - the problem of evaluating the quality of reading outcomes.

1.4 Evaluating the Quality of Reading Outcomes

1.4.1 A Definition of Quality

Stern argues that "language teaching research, in the first instance, is educational research, and the principles and procedures of research in education and the behavioural sciences are applicable." (Stern, 1983, p.63). It seems to me, consequently, that foreign language reading research can benefit from exploring the ways in which researchers working in the educational field have dealt with the problem of qualitative assessment.

However, since the way in which the term "qualitative" is used varies, some terminological clarification will be convenient at this stage.

Grotjahn points out that "the term 'qualitative' is used in at least two different senses," and that "it frequently remains unclear which sense is being used in the discussion." (Grotjahn, 1987, p.57).

The case of the term "qualitative" is just an instance of a phenomenon commonly observable in research carried out in education and the behavioural sciences at large, for, as Perry puts it: "...it is inevitable in so complex a field, drawing on research findings suddenly emerging in diverse sources, that terms and concepts should be unstable." (Perry, 1983).

Let's go back to the distinction established by Grotjahn:

With regard to concepts and data used to grasp a certain domain of reality, the term "qualitative" designates either a certain type of concept or the level of measurement of the data to be collected. In general, classificatory concepts and thus, also, data measured on a nominal scale are considered to be qualitative, whereas metrical concepts or data measured at least on an interval scale are regarded as quantitative. (op.cit., p. 57).

To illustrate this point, consider the following definition of

"nominal scales" given by Smith (1975,p.42).

Nominal Scales refer to qualitative observables which are only arbitrarily classifiable in terms of same or different, (i.e., male-female), equivalent or non-equivalent (i.e., family, non-family), or either-or terms (i.e., similar or non-similar attitudes).

The other sense in which the term is used, according to Grotjahn, is the following. He states:

- On the other hand the term "qualitative" is often used in a much wider sense to designate an entire research paradigm.... It then refers simultaneously to the manner of data collection, of theory construction and of data analysis as well as to the social and philosophical orientation of the investigation in question.

(op.cit., p.58)

He then notices that when used in this wider sense, it is frequently used in opposition to the term "quantitative." He finds that the opposition "qualitative" v. "quantitative" has been characterized in several ways by different authors. (Ibid,pp.58-9). Some of such ways are listed below:

- Naturalistic (unobtrusive) and uncontrolled observation v. obtrusive and controlled measurement.
- subjective v. objective.
- discovery oriented, exploratory, descriptive and inductive v. verification-oriented, confirmatory, inferential and hypothetico-deductive.
- ungeneralizable, single case studies v. generalizable, multiple case studies.
- phenomenology... v. logical positivism.
- valid, 'rich', and 'deep' data v. reliable, 'hard' and replicable data.
- interpretative v. statistical.

It seems to me that the following statement taken from the foreword to the book Understanding Student Learning, illustrates this dichotomy: "... (the authors) have combined both quantitative and qualitative modes of inquiry... they have made a productive synthesis through interweaving of quantitative analysis of questionnaires with the qualitative assessment of interviews." (Perry, 1983).

In view of the differences mentioned above, it will be necessary to define the way in which the term "qualitative" will be understood in the present study.

As already mentioned, the SOLO Taxonomy (Structure of the Observed Learning Outcome) will provide the basic tool for the analysis of the data in this study. So, the term "qualitative" will be used in the sense in which the authors of the Taxonomy have used it when addressing the issue of "quality" and "quantity" in learning.

They state, for instance, that "the educator needs to evaluate (as indeed does the student) how much has been learned and how well it has been learned." (Biggs and Collis, 1982,p.3). Then, following Marton (1976), they continue: "evaluation involves both these quantitative (how much) and qualitative (how well) aspects."

The distinction is made, therefore, not in terms of the scale of measurement used (nominal = qualitative v. other = quantitative), but rather in terms of the manner of data analysis.

A further clarification regarding the way in which "qualitative" and "quantitative" will be used in this study is still in order. If "qualitative" will basically apply to the "how well" aspect of the dichotomy - as established by Marton, and Biggs and Collis -, "quantitative" will not refer to "how much", but rather, to "how many". That is, the analysis of the subjects' reading outcomes will be carried out "qualitatively" first (by using the SOLO Taxonomy, as described in later chapters of the thesis); and then, the results will be submitted to "quantitative" statistical analyses, to examine "how many" subjects fall within the contrasting groups of the experiment, and to determine if the differences between the groups are statistically significant.

Let us now consider, in the following sections, some of the ways in which the problem of qualitative assessment has been tackled in education.

1.4.2 Evaluating Quality in Education

Perhaps the best-known and influential instrument for defining quality in educational assessment is the Bloom Taxonomy. (Bloom, 1956). A short but clear account of this taxonomy can be found in Martuza (1977, pp.187-97). The definitions of its six main categories that follow have been taken from that source.

A knowledge level (or memory level) item measures the student's ability to recall or recognize information (e.g., terminology, facts, principles, problem-solving strategies) in essentially the same form as it was presented to him. The essential student behaviour elicited by questions at this level is remembering. At all higher levels of the taxonomy, remembering is only one of several behaviours required to successfully answer the items presented to the student....

Comprehension level items measure understanding at the most rudimentary level. Items at this level measure the student's ability to use previously acquired information to solve a novel problem. A key feature of measurement at this level is that the item or the context in which it is asked is structured in such a way that the student is made aware of the information required to solve the problem....

An application item, like the comprehension item, requires the student to use previously acquired information in solving a novel problem. Unlike the comprehension item, however, neither the question nor the context in which it is asked helps the student decide what previously acquired information must be used to solve the problem. Questions at this level are aimed at determining whether the student is able to select the appropriate knowledge as well as use it correctly in solving the novel problem....

An analysis level item may require the student to (a) identify a logical error in a communication (e.g., a contradiction, an error in deduction, an erroneous casual inference) or (b) identify, classify, and/or recognize the relationships among the elements (i.e., facts, assumptions, hypotheses, conclusions, ideas, etc.) in a communication. Items at this level usually assume specific training in the logical process to be used....

A synthesis level item requires the student to create (a) a unique verbal or nonverbal communication (e.g., a story about a personal experience, an extemporaneous speech, a poster showing the evils of drug abuse), or (b) a plan or procedure for accomplishing a particular task (e.g., designing a simple experiment to test an interesting hypothesis, formulating a plan for erecting a tent)....

An evaluation level item requires the student to judge the value of ideas, people, products, methods, etc. for a specific purpose and state valid reasons for his judgement.

Biggs and Collis point out that the taxonomy "is really intended to guide the selection of items for a test rather than to evaluate the quality of a student's response to a particular item." (Biggs and Collis, 1982,p.13). It represents, as Dahlgren notices, a logical analysis of the content of learning, but it does "not derive from studies of the different outcomes arrived at for a given subject-matter." (Dahlgren, 1984,p.28).

On the other hand, there have been some attempts to design tools to evaluate the quality of the student's actual performance. I will now refer to two that are closely related to the SOLO Taxonomy, which is the one used in this study.

The authors of the SOLO Taxonomy cite the work of Schroder, Driver and Streufert (1967). Their classification of levels of "information processing" was developed within the field of experimental social psychology, but they claimed it could also be used to mark essays. The classification includes four levels of "conceptual structure," i.e., "the way an individual receives, stores, processes, and transmits information":

1. Unidimensional, characterized by minimal conflict, rapid closure, and categorical judgement.
2. At least two dimensions are attended to, which may lead to conflicting and inconsistent judgements.
3. Uses subordinate rules to relate inconsistencies and resolve conflict, choice which is less determined by external forces.

4. Theoretically oriented, able to generate own rules to cover all cases, including hypothetical ones.
(Biggs and Collis, 1982, p.13).

Biggs and Collis point out that their own scale extends more widely than that developed by Schroder, Driver and Streufert, and that it is probably more easily applied to a wider range of educational situations. (op.cit.p.14).

The work of Peel (1971) is also a close antecedent of the SOLO Taxonomy. Entwistle points out that Peel explored "systematically the types of thinking exhibited by pupils in a variety of content areas," and notes that the pupils' responses usually fell within one of three main types:

- Mentioning - tautological, partial and inconsistent responses.
- Describing - a mainly correct listing of aspects of the passage.
- Explaining - responses which used additional related ideas to interpret the meaning of the passage.

(Entwistle, 1985, p.6)

Also, a more developed form of explanation was sometimes found. Such responses were characterised as "a capacity to combine more than one piece of evidence with outside ideas and to evoke cause and effect." (Ibid).

The relationship of these two classifications to the SOLO Taxonomy will become clear in the next section, where the latter will be introduced. In that section, attention will also be given to other three models that may provide alternative frame works for evaluating the quality of reading outcomes. (The Brunel Model, the Gothenburg model, and the Kintsch and van Dijk model).

1.4.3 Evaluating Reading Outcomes

In this section, I will first describe the way in which two research groups have approached the task of analysing university students' reading performance, in Brunel and Gothenburg universities, respectively. Afterwards, I will introduce the SOLO Taxonomy, and explain the reasons for choosing it rather than the other two models. Then, I will briefly refer to the Kintsch and van Dijk model, which has been influential in the study of summary writing, the basic task the subjects in this study were required to carry out. I will, in this case too, point out why the SOLO Taxonomy was considered a more convenient research tool for this study. ✓

1.4.3.1 The Brunel Model

Substantial research on the topic of learning to learn, with particular reference to reading, has been carried out in Brunel University since 1962. (Cf., Thomas, Augstein and Farnes, 1973). As can be expected, such a vast project has covered a wide range of related topics, as well as yielded important and varied results. So, the following paragraphs will serve the purpose of presenting a condensed summary of the main aspects of the project.

An essential step in improving Reading-for-Learning skills is to increase the reader's awareness of what he is doing as he reads. For the unpractised, self-observation during reading is immensely difficult. It is necessary to acquire the skill of dividing one's attention between the meaning of the text and how one is attributing that meaning to it. This skill in self-observation derives from a language which is used to describe and analyse one's reading experience and behaviour. In these studies, our approach has been to offer a Model which provides such a language and also to offer tools which reflect back different aspects of his process to the reader. These tools can be considered as consciousness-raising devices.

The Reading Recorder provides a record of how a text is read. This record offers the reader the opportunity to introspect the process in which the eyes move over the page. The Flow Diagram Technique provides an opportunity for the reader to display the structure of meanings he is attributing to the text. This increases his awareness of how he develops and uses meanings. The Kelly Repertory Grid Technique reveals the dimensions of thought and feeling. We have used it to help the reader explore his ideas about learning and about different types of text. The Purpose Taxonomy is also essentially an awareness creating device; the Reading-for-Learning Exercises were based on this Taxonomy. During the project these tools were further refined and other tools were developed and tested.

(Thomas and Augstein, 1976c, p1).

As for the model of the reading process that has been used, it is stated below:

Reading is considered as an active generative process, which operates at several levels simultaneously. The major components of the process at any given level are considered to be: ... PURPOSE, STRATEGY... OUTCOME...REVIEW....

Although purpose partially determines strategy and this influences outcome they must be seen as co-existing. The purpose is never finally clarified until the final outcome is achieved.

As a person reads, he is predicting meanings that will be contained in the words on the page. The Strategies and Tactics of Reading Behaviour relate to Prior Knowledge and Skills and to Reading Purposes. Reading Behaviour also relates to the Structures of Meaning which a person generates. These structures are provoked by the words on the page, though the meaning itself is not on the page. The primary outcomes of reading are the changes which take place in what we think, know and feel. These changes must serve as the ultimate reference of whether there has been any reading outcome. Any oral and written Reading Outcome has structure, which can be observed and thus directly measured. This plays a crucial role as the criteria against which the reading process can be assessed. Thus: improvement of Reading Competence depends on monitoring the matrix of Purposes, Strategies and Outcomes and Reviewing their effectiveness. This review can then be used to improve subsequent attempts at reading which should in turn be evaluated, creating a progressive and cumulative process of change.

(op.cit., pp.2-4).

Finally, regarding the research approach used, "the conversational approach," Augstein and Thomas point out:

The conversational approach adopted in our investigations accepts the reader/learner as full participant in the descriptive and interpretative enterprise. It uses the reader/learner's unique position as observer of those internal events whereby meaning is attributed to a text, and admits the personal validity of the reader's attempts to explain such personal processes. The tutor with the technology at his or her disposal can observe and record external aspects of the reader's behaviour. Such records can then be used to help the reader to introspect and describe personal reading events. The advantage of the conversational approach lies in involving the reader in the observation and explanation.

(Augstein and Thomas, 1984, p.252).

1.4.3.2 The Gothenburg Model

This model is also the product of a long series of research projects carried out at the University of Gothenburg, Sweden, since the 1970's. The researchers have been concerned with the quality of learning of university students in general, but have also paid specific attention to the process of reading. As the Brunel researchers, they have also worked with students reading in their mother tongue. Their work has already gained recognition amongst educators working in the area (cf., for example, Biggs and Collis, 1982; Entwistle, 1981; Perry, 1983; Pollitt, et al., 1985). It seems to me, that the study of their findings and insights into the reading process as well as their research methodology can also be profitable in the case of the foreign language reading research. I will now mention some aspects of these studies that seem to me important in this context.

1. Their main concern has been the qualitative - rather than the quantitative - analysis of the process and outcome of learning. (Cf., for example, Marton and Säljö, 1976a).

2. The studies have covered a wide range of aspects of the learning process. For instance, they have studied possible relationships between the way in which students read and factors such as:

- the content of the particular text and task. (e.g., Marton and Säljö, 1976a).
- The nature of the reading task. (e.g., Marton and Säljö, 1976b).
- the learner's conception of the task. (Ibid).
- type of motivation (intrinsic or extrinsic). (Fransson 1977; and 1984).
- level of anxiety while reading. (Fransson, 1984).
- students' approach to their normal studies (Svensson, 1977; and 1984).

3. As for the methodology used, the researchers have adopted a phenomenological approach - called phenomenography by Marton (cf., Marton, 1981; and Entwistle and Marton, 1984) - which takes into account the students' perceptions of their learning experience, as reported in interviews. The interviews may be followed by experiments testing the results of the interviews. Thus, the results can be said to be less affected by the preconceptions of the researchers.

4. As for the reading outcomes, Entwistle (1981,p.76) summarizes the four types of response identified as the basic framework in the Gothenburg studies in the following terms:

A. Conclusion-orientated, detailed.

The student summarizes the author's main argument, shows how evidence is used to support the argument, and explains the thoughts and reflections used to reach personal understanding of the argument.

B. Conclusion-orientated, mentioning.

Again there is an adequate summary of the main argument, but the use of evidence or personal experience to support that argument is not made clear.

C. Description, detailed.

The student gives an adequate list of the main points presented in the article, but fails to show how these are developed into an argument.

D. Description, mentioning.

A few isolated points are made, some relevant, others irrelevant. At the bottom end of this category an impression of confusion and misunderstanding is given by the student's comments.

1.4.3.3 The SOLO Model

The SOLO Taxonomy (Structure of the Observed Learning Outcome) was developed by Biggs and Collis after studying the organisation of responses from many hundreds of students from elementary to college levels, to evaluate learning quality in a wide variety of subjects such as history, geography, mathematics and reading. (Cf., Biggs and Collis, 1982).

The following is a clear and brief description of the taxonomy made by Dahlgren (1984, p.28), and can be useful as an introduction to it. This description can also make clear its relationship with the classifications of Schroeder, Driver and Streufert, and Peel, mentioned in Section 1.4.2 above. Dahlgren states that the SOLO Taxonomy . . . is

an attempt at empirical classification of levels of outcome in a form which has wide applicability. The theoretical basis of Biggs and Collis' taxonomy derives in part from the stages in cognitive development described by Piaget and in part from theories of information processing...Breaking away from Piaget's use of stages to describe the developmental level of an individual child, Biggs and Collis seek to describe the range of answers given to a specific question...They assume that such levels have a general reality, irrespective of content and question form, and describe five categories as follows, with increasing levels of sophistication..

1. Pre-structural. In relationship to the prerequisites given in the question, the answers are denying, tautological, and transductive. Bound to specifics.

2. Uni-structural. The answers contain "generalizations" only in terms of one aspect. ✓

3. Multi-structural. The answers reveal generalizations only in terms of a few limited and independent aspects.

4. Relational. Characterized by induction, and generalizations within a given or experienced context using related aspects.

5. Extended Abstract. Deduction and induction. Generalizations to situations not experienced or given in the prerequisites of a question.

It may be added here that intermediate - transitional - categories can be identified, thus complementing this basic set of categories.

Now, as indicated at the beginning of this section, I would like to point out why the SOLO Taxonomy was chosen as the basic research instrument to be used in the present study, rather than those provided by the Brunel or Gothenburg investigations.

Although the three groups of researchers can define attainment qualitatively, it seems to me that the SOLO model provides a more complete taxonomy of possible outcomes than the Gothenburg basic categories, while at the same time - because in this respect its intent is different - it provides a more specific set of possible outcomes than in the case of the more open Brunel approach.

More important, though, is the fact that both the Brunel and Gothenburg research methodologies require constant interaction with the subjects, whereas the use of the SOLO Taxonomy does not. This made it particularly useful for the purposes of this study, where constant participation of the rather large number of subjects involved would have been difficult to secure, and would have required a much longer time than was actually available.

Before closing this section, I would like to mention another research tool that has been influential in the study of summarization processes and outcomes, and to point out as well why the SOLO ✓

Taxonomy was chosen for the present study. The tool in question is the Kintsch and van Dijk model of "text comprehension and production." (Kintsch and van Dijk, 1978). This model developed from a different academic perspective (that of discourse studies) from the one which has been mentioned above and within which the interest of the present study arose, namely, the wider educational concern of qualitative assessment.

Briefly stated, as Winograd (1984, p.405) puts it, the model

attempts to show how individual propositions in the text are transformed and condensed into the gist. The essential components of this model are the reader's schema, the microstructure and the macrostructure of the text, and a set of macro-rules for producing summaries. According to this model, readers progress through a text reducing and organizing its microstructure into a macrostructure through the application of a series of transformations known as macro-rules.

This model has provided the basis for studies dealing with text summarization both in LI (cf., for example, Brown and Day, 1983; Singer, 1982; and Winograd, 1984), and in a foreign language (cf., for example, Kozminsky and Graetz, 1986; and at the National University of Mexico: Cao-Romero, 1983; and Ryan, 1987).

✓ As for the present study, it has already been mentioned that the subjects were asked to perform two very closely related tasks, that is, to write a summary of the text read, and to make their own comments afterwards. The Kintsch and van Dijk model may have been used for the analysis of the summaries, but I needed a tool that would facilitate an integrated analysis of summary and comments. The SOLO Taxonomy, given its versatility offered a better option in this respect. Also, besides its suitability for the qualitative analysis of summaries and comments, the SOLO Taxonomy provides a scale that lends itself to the quantitative analysis intended.

In the next chapter, a more complete discussion of the way in which the SOLO Taxonomy can be applied to the analysis of reading outcomes as outlined above will be presented, and an exploratory study will be reported.

CHAPTER TWOThe Solo Taxonomy And Reading Outcomes

2.1 Introduction

In this section we will introduce the SOLO Taxonomy as an instrument devised to provide a criterion to evaluate the quality of learning outcomes. In the next section the SOLO Taxonomy will be described in detail. Then, a discussion will follow of how it can be applied to analysing reading outcomes of the type of those elicited in the present study. An exploratory study to determine the viability of using it to analyse summaries produced after reading a text in both a foreign language (English) and the subject's mother tongue (Spanish), will also be reported.

All quotations from Biggs and Collis's SOLO Taxonomy are taken from the 1982 edition of their Evaluating the Quality of Learning book, unless otherwise specified.

Let us now turn our attention to the type of assessment that the SOLO Taxonomy lends itself to.

The quantitative aspects of assessment have received attention in education and psychology for quite some time, and a solid body of research and statistical techniques has already been developed in this area.

However, the qualitative aspects have not been so widely studied. It could be said that the SOLO Taxonomy was developed as a response to the needs felt in this other area. More specifically, the aim of its authors was to provide a "criterion-referenced measure of the quality of learning." (p.8).

It may be convenient at this stage to briefly define what is meant by a "criterion-referenced" measure. The term is defined in evaluation in opposition to "norm-referenced" testing. Ingram states

the contrast between the two types of testing thus: "norm-referenced testing compares the behaviour of an individual with the behaviour of others, while criterion-referenced testing describes the behaviour of an individual with reference to externally predetermined and specified objectives." (Ingram, 1977, p.26). In other words, in criterion-referenced testing, the main interest is placed in finding out "whether or not the individual student knows something rather specific that he is supposed to know, or can perform something rather specific that he is supposed to be able to perform". (Ibid). She goes on to point out, regarding the specification of the criterion, that it can be done in terms of content (e.g., "does the learner know, or does he not know how to construct a question in English using 'do'?"), or in terms of degree of skill or levels of performance (e.g., "does the learner possess the skills which would allow him to function adequately as a tourist or as a student of current affairs?"), and that "this kind of performance criterion is much harder to define." (Op.cit., p.27).

It is for this purpose that the SOLO Taxonomy can be of great help. By providing a tool for evaluating the quality of learning outcomes - by analysing their structure - it can also serve the purpose of specifying criteria against which to determine if the learner has achieved the desired standard qualitatively speaking, or not. In other words, one may be able to assess how well the prescribed standard (specified in terms of content or skills) has been met. It is in this sense that the taxonomy can provide a "criterion-referenced measure of the quality of learning."

Let us now examine the taxonomy in more detail.

2.2 Description of the SOLO Taxonomy

I shall now expand the information about the SOLO Taxonomy provided in Section 1.4.3 above, with information about other relevant aspects of the taxonomy. I shall refer, for this purpose, to the authors' description of its basic features. (Table 2.1).

SOLO description	1 Capacity	2 Relating operation	3 Consistency and closure	4 Response Structure Cue Response
Extended Abstract	<i>Maximal:</i> cue + relevant data + interrelations + hypotheses	Deduction and induction. Can generalize in situations not experienced	Inconsistencies resolved. No felt need to give closed decisions—conclusions held open, or qualified to allow logically possible alternatives. (R_1 , R_2 , or R_3)	
Relational	<i>High:</i> cue + relevant data + interrelations	Induction. Can generalize within given or experienced context using related aspects	No inconsistency within the given system, but since closure is unique so inconsistencies may occur when he goes outside the system	
Multistructural	<i>Medium:</i> cue + isolated relevant data	Can "generalize" only in terms of a few limited and independent aspects	Although has a feeling for consistency can be inconsistent because closes too soon on basis of isolated fixations on data, and so can come to different conclusions with same data	
Unistructural	<i>Low:</i> cue + one relevant datum	Can "generalize" only in terms of one aspect	No felt need for consistency, thus closes too quickly jumps to conclusions on one aspect, and so can be very inconsistent	
Prestructural	<i>Minimal:</i> cue and response confused	Denial, tautology, transduction. Bound to specifics	No felt need for consistency. Closes without even seeing the problem	

*Kinds of data used: X = irrelevant or inappropriate; ● = related and given in display; ○ = related and hypothetical, not given.

Table 2.1 Basic Features of the SOLO Taxonomy.

(based on Biggs and Collis, pp.24-5, Table 2.1).

Before looking at what the authors have to say about the different categories in Table 2.1, it will be convenient to reproduce the answers they analyse in their discussion. (Figure 2.1).

1. "Dunno."
2. "Because it rains more on the coastal side."
3. "Because when we go to our cabin that's right on the coast, it's always wetter there than on the road crossing the mountain that gets us there. Never fails, my Pop says. I reckon we ought to move; like get us a cabin for hunting which is better'n fishing anyway. Besides I hate rain."
4. "Because the sea breezes hit the coastal side of the mountain first."
5. "Cos air from the sea gets kinda damp, like fog and that. It settles on the coast first and so it rains and all the wetness falls on the coast and there's none left for the other side of the mountain."
6. "Because the prevailing winds are from the sea which is why you call them sea breezes. They pick up moisture from the sea and as they meet the mountain they're forced up and get colder because it's colder the higher you get from sea level. This makes the moisture condense which forms rain on the side going up. By the time the winds cross the mountain they are dry."
7. "This is likely to be true only if the prevailing winds are from the sea. When this is so, they pick up the water vapour evaporated from the sea which is carried to the mountain slopes where the damp air mass rises and cools. Cooling causes the water vapour to condense, and being heavier than air, the water droplets deposit as rain. Not only is the wind now drier, it is possible that it is carried up the mountain further where it is compressed, which warms it like a bicycle pump gets warm. It is therefore less saturated than before for two reasons. The effect is like the chinooks experienced on the eastern slopes of the Rockies in Canada in winter. If there was no mountain, there would likely be no difference between the coast and inland. It all depends on the land features and the prevailing wind and temperature conditions. If these differed, then the energy exchanges would be different, resulting in quite a different pattern."

Figure 2.1 Answers to the question "Why is the side of a mountain that faces the coast usually wetter than the side facing the interior?" (Used by the authors as an example of the way in which the SOLO Taxonomy can be used in qualitative analyses). (Biggs and Collis, pp.4-5).

We are now in a better position to follow the authors' discussion of those features of the SOLO Taxonomy in Table 2.1.

Regarding the feature "capacity," it is to be understood as "the amount of working memory, or attention span, that the different levels of SOLO require." (Biggs and Collis, p.26).

To illustrate how this can be applied to the analysis of learning outcomes, they discuss the coastal rain examples. They point out that the first three examples are prestructural: in Example 1, the student does not even have to remember the question itself, whereas at least this is necessary for Examples 2 and 3:

A prestructural response makes the least demands: One "bit," the cue (i.e., the question) and response are confused into one global undifferentiated unit. The simplest relevant response requires separation between the cue and one relevant aspect to form the response. In other words, in order to answer the question adequately, the student must, at the least, bear the question in mind while he is answering, and then try to relate the question and the answer with at least one logical operation. Thus the response in Example 4 is unistructural; it relates one relevant datum to the question. The capacity required is that which enables the handling of two "bits" (cue plus one relevant datum). A multi-structural response (Example 5) involves two or more concepts or data; a relational response (Example 6) goes further and interrelates the concepts, so that in this case we have the cue plus most or all relevant data plus their interrelationships. The extended abstract (Example 7) goes even further than that; the student here needs not only to encode the given information, but to comprehend its relevance to overriding abstract principles, from which he can deduce a hypothesis and apply it to a situation that is not given: All this obviously requires a larger capacity of working memory than the other SOLO levels. (Ibid.).

The next column in Table 2.1 specifies the "relating operations" for each level of the taxonomy. This feature

"refers to the way in which the cue and the response interrelate." (Ibid). After pointing out that "in the case of prestructural responses there is no logical interrelation," the authors define the three types of confusion they have identified at this level (p.27). The first one is "denial," in "which the student refuses to become seriously engaged in the task." (Cf. Example 1). The second one, "tautology," is simply a restatement of the question. (Cf. Example 2). In the case of "transduction," there are "attempts to differentiate a relevant response but [the student] slips up because she does not form an adequate logical basis for her selection of a response." (Cf. Example 3).

Regarding the way in which "induction" and "deduction" are to be understood in this context, the authors' discussion of the coastal rain examples will again be useful.

Induction involves correctly drawing a general conclusion from particular instances; or in the present context, it means relating a particular aspect or point given in the data to a conclusion. A unistuctural response, then, involves ascribing rain to only one of the relevant features. A multistuctural response faithfully marshalls several of the relevant features but fails to link them up. The multistuctural response typically contains "...and...and so...and also...." The relational response gives an overall concept or principle that accounts for the various isolated data that the multistuctural response contains, but sticks within the data and concepts already taught about the formation of rain.

The extended abstract response goes beyond induction on the basis of the data and introduces true logical deduction: "Yes, the phenomena is due to the principles of heat exchange. Now, given the details that are described, the rain will precipitate in the prescribed area. But under other conditions the outcome would probably be different. We see the same principle operating in the quite different context of the Canadian Rockies...." There are several features about the extended abstract response that should be noted: (a) the introduction of an abstract principle (principle of heat exchange) which was not given directly in the data; (b) the deduction from the principle that certain events would follow, and the testing of this deduction against the data; (c) the introduction of an analogy (the Rockies in Canada) that was compatible with that principle but not given in the data and consequently; (d) the outcome may be indeterminate (i.e., there is an absence of closure-events might have been different under different circumstances).

(Ibid.).

The features "consistency" and "closure" in Table 2.1 (Column "3"), refer to two opposing needs experienced by the student. On the one hand, he feels the need to "come to a conclusion of some kind (to close)." On the other, he feels the need to "make consistent conclusions so that there is no contradiction... between the conclusion and the data." (Ibid.).

Let us now look at the way in which the authors see the relationship between these two features in the examples under consideration. They start with prestructural responses, noticing that they are marked by

very high closure and very low consistency: In fact, the student at this level closes by simply saying "I dunno," repeating the question, or by transducing across some irrelevancy. The student who gives a unistructural response seizes upon the first relevant dimension that comes to mind, but at least it is relevant. Unistructural responses can all be equally correct, but quite inconsistent with each other, rather like the blind men describing the elephant: The elephant was described as a piece of rope, a wall, a snake, a tree trunk, all correctly given the perspective of each man.

In multistructural responses, closure is determined when more aspects are perceived, but since these aspects are not interrelated, inconsistency may result. Two responses at this level may utilize the same amount of data but come to quite different conclusions.

The student responding relationally waits until he sees the aspects and then interrelates them to make a coherent whole. He will come up with a definite answer (closure), possibly an excellent answer for that context, but it will not do for other contexts (i.e., an over-generalization may be made). The relational response is still tied to concrete experience: It leaves room for inconsistency across contexts. For example, the factors that produce coastal rain may not always be present. The extended abstract response on the other hand, sets out the principles and heavily qualifies their application to a particular situation (i.e., consistency is maximal, and, as a result, the student may feel it appropriate to leave the question relatively open).

(p.28)

The last column in Table 2.1, "response structure," is a diagrammatic representation of the characteristics of each level of the taxonomy.

The authors point out that there are three types of data that students may use in response to the question:

- irrelevant data (represented by X).
- relevant data that are contained in the original display of information (e.g., lesson), (represented by ●).
- relevant data and principles that are not given, though they are often implicit in the data, or hypothetical but relevant (represented by ○).

(Ibid.).

They then describe what happens at each level of response in this respect.

At the prestructural level there may be an attempt to link the cue with the response by an irrelevant feature. The unistructural response takes one relevant datum or feature to link the cue and response; the multistructural response takes several. The relational response ties up the relevant data in a conceptual scheme. The extended abstract response takes up all the relevant data and their interrelations and subsumes them under a hypothetical abstract structure that can enable deductions to apply to instances or data which were not included in the original display. Consequently, the student giving an extended abstract response can entertain alternative outcomes: He is not forced, as are the others, to come to a definite closure or conclusion.

(Ibid.).

Another aspect of the taxonomy that I want to mention here, is the fact that sometimes it is convenient - and possible - to expand this basic set of categories by including intermediate, "transitional," ones. As Biggs and Collis point out (p.29), there may be some responses that will not fit neatly into the basic set. "Typically," they say regarding transitional responses, they "carry more information than is usual in the level the student is emerging from, but he is forced to give up before reaching the complexity of structure that is expected at the next SOLO level." There are, then, four transitional categories: from

prestructural to unistructural (which the authors refer to as 1A), from unistructural to multistructural (2A), from multistructural to relational (3A), and from relational to extended abstract (4A).

The following classification illustrates how the transitional categories can be linked to those in the basic set. This was the basic frame of reference which the authors used to analyse responses to history questions. (Figure 2.2).

1. Prestructural. Student avoids the question (denial), repeats the question (tautology), a firm closure based on transduction.
 - 1A. Transitional. Student attempts to answer the question but only partially grasps a significant point.
2. Unistructural. An answer is based on only one relevant aspect of the presented evidence so that the conclusion is limited and likely dogmatic.
 - 2A. Transitional. An attempt to handle two aspects of the evidence is made, but they may be inconsistent and hence no firm conclusion is reached.
3. Multistructural. Several consistent aspects of the data are selected, but any inconsistencies or conflicts are ignored or discounted so that a firm conclusion is reached.
 - 3A. Transitional. Any inconsistencies are noted: Several aspects are recognizable but the student is unable to reconcile them.
4. Relational. Most or all of the evidence is accepted, and attempts are made to reconcile. Conflicting data are placed into a system that accounts for the given context.
 - 4A. Transitional. There is a hint that closure, or a firm conclusion, is not inevitable. There is a suggestion that a relating principle might account for the situation, but this is not spelled out.
5. Extended abstract. There is recognition that the given example is an instance of a more general case. Hypotheses about not given examples are entertained, and the conclusions are held open.

Figure 2.2 The Complete SOLO Taxonomy, as Applied to History Items by its Authors. (Biggs and Collis, p.36).

I would like to mention at this stage that in the exploratory study reported later in this chapter, no transitional categories were used. However, thanks to that study, it became clear that the marking scheme to be used to analyse the performance of the

subjects of the experiment would benefit from their inclusion. Accordingly, as will be seen in Chapter 3, the complete set of categories in the taxonomy was used in the full experiment.

In the next section, we shall take a closer look at how the SOLO Taxonomy can be used as a research tool in a study on reading comprehension.

2.3 An Application of the SOLO Taxonomy to the Analysis of Reading Outcomes: An Exploratory Study in LI and FL

2.3.1 Background

Having decided to use the SOLO Taxonomy to evaluate the subjects' reading outcomes in this investigation, the next stage was to determine how to do so in the most adequate manner.

Previous studies reported by its authors in which the taxonomy was used within the areas of reading, and foreign languages, were not directly related to the present study.

In the case of reading, the studies reported were carried out to analyse problems in the teaching of early reading, and in LI only. They concentrated, more specifically, on the study of "word attack skills" used by children in elementary schools. (Cf., Biggs and Collis, pp.104-8).

In the case of foreign languages, the authors give examples of how the SOLO Taxonomy may be applied to two different tasks: translation of vocabulary items, and derivation of a grammatical rule from particular instances of its use (op.cit., pp.145-160).

Since the authors did not report any research in the specific area of my concern, the present study would be, as far as I know,



the first one to analyse summaries written after having read a text in both, the subjects' mother tongue and in a foreign language. So, a small scale experiment was devised to explore the viability of using the taxonomy as I intended. A report of it follows.

2.3.2 Analysis of the Materials Used from a SOLO Perspective

In this exploratory study, the subjects were asked to perform the same task in a foreign language (English) and their mother tongue (Spanish). Two tests were designed for this purpose:

Test A: In this test, the subjects were asked to state in their own words the topic of the text, allowing them to make their own comments on the material read. The text was in English. (Please see Appendix A).

Test B: Same task as in Test A. Text in Spanish. (Please see Appendix B).

The question asked is similar to one used by Entwistle in a test with basically the same purposes as mine:

Write down what you have learned from the article. Imagine you were going to describe what the article is about to someone who had not read it. What would you say?

(Entwistle, 1981, p.61)

The instructions were given in Spanish, and the subjects requested to answer in Spanish.

Besides controlling for reading task by asking the subjects to do the same in both languages, I controlled for content, nature and degree of difficulty of text. This was done by using a translated text. I preferred to use a translated text, rather than "matched" texts, because I agree with Meara when he states

that "no two sets of matched passages are ever matched identically on all variables" (Meara, 1984, p.99), and with Urquhart: "... to find suitable pairs of already existing texts, requires perhaps too much searching, not to mention luck, to be practical." (Urquhart, 1984,p.162). I translated the original into Spanish myself, with the help and revision of two other native speakers of Spanish. As far as subject matter is concerned, the general topic "ecological problems of Mexico City" was considered suitable for readers with different academic backgrounds.

From a formal point of view, it should be noted that the text is a complete unit. That is, it is a whole newspaper article, not an extract from a wider context.

Regarding its length, since it is not too short it is not unrealistic to ask the subjects to summarize it. Yet, it is not too long to be reasonably managed in a test situation.

Finally, it seemed to me that the way in which it is structured would facilitate its initial analysis in terms of the SOLO Taxonomy. Such analysis provides a basic framework against which to study the subjects' responses. So, it will be convenient at this stage to take into consideration some concepts expressed by Biggs and Collis regarding task analysis. They say, for example, that

...it is necessary to define what the minimal features are such that any one of them could be a unistructural response and several of them could make up a multistructural response. Some sort of relating concept then needs to be identified in the passage for a relational response. An extended abstract pro forma is often unpredictable, since different respondents might construct different hypotheses, but at least it is clear when this does happen. In brief, there are inherent in the data a series of whats that need to be identified; and the teacher has to decide what is a what.

(Biggs and Collis, pp.165-6).

In their discussion they include what seemed to me useful guidance for identifying at a more practical level the relevant components of a task.

They say, for example, that the components in a unistructural response would consist of "all the basic facts, details, and concepts that are given in the original presentation. Any one of these would constitute a unistructural response (if it addressed the question or requirement appropriately)." (Ibid, p.169).

They also point out that several of such basic facts, details or concepts could make up a multistructural response. Then, for a relational response, an integrating concept should be identified in the passage. "There might be several possibilities. These would correspond. . .to the list of components in the [relational] level." (Ibid). The components of an extended abstract response cannot be prescribed in advance.

Table 2.2 is an attempt to apply what has been stated so far about the SOLO Taxonomy to the practical task of establishing expected responses in this exploratory study. As it can be seen, the first two columns comprise relevant information about the way in which different levels can be established according to the authors of the taxonomy. The third one is an application of such information to the specification of expected responses in this study. As it stands, then, the table was used to analyse and classify as objectively as possible the subjects' responses.

LEVEL OF RESPONSE	MAIN CHARACTERISTICS OF RESPONSES DEPENDING ON LEVEL	RESPONSES THAT MAY BE EXPECTED IN THE STUDY: MAIN CHARACTERISTICS
P r e s t r u c t u r a l	The basic information in the passage has not been re-covered correctly.	Responses at this level are very unlikely. However, a summary could be included here if - for example - the causes of the different illnesses mentioned in the passage are not properly identified as such; and/or, what the various people and institutions mentioned in the passage have said or done in relation to pollution in the city is misunderstood.
U n i s t r u c t u r a l	Any of the basic facts, details and concepts that are given in the original passage would constitute a uni-structural response (if addressed the requirement appropriately).	Summaries to be included at this level would be those that - even if correctly - would deal with just one aspect of the original passage. For example, a summary that states correctly what has been found regarding lead poisoning in Mexico City, but does not go beyond that, would fall within this category.
M u l t i s t r u c t u r a l	Several of the basic facts, details and concepts that are given in the original passage could make up a multi-structural response (if addressed the requirement appropriately).	Summaries in this category would be basically an enumeration of data, without establishing clearly the way in which such data are interrelated. An example would be a summary en-listing all the illnesses that are mentioned in the article, without at least grouping them by related cause. These summaries, though conveying correct information, fail to give a coherent view of the whole original passage.

R e l a t i o n a l	An integrating concept in the original passage is recovered. (There may be several possibilities).	Summaries in this category should establish clearly the way in which the data in the original are inter-related. For example, they could at least establish which groups of illnesses are caused by which types of pollutants. Other relating concepts in the text could be: <ul style="list-style-type: none">- the concern shared by various groups and institutions about the problem.- the limited action taken by the government to remedy the situation.- the lack of reliable official information on the subject.
E A x b t s e t n r d a e c d t	Data, principles, concepts, hypotheses, that are not given in the original passage, but which are relevant are included in the response.	The form an extended abstract response will take is by definition unpredictable, since the subject goes beyond the text, reflects upon the material read, takes a critical stance, and makes a personal contribution.

Table 2.2 Examples of responses that may be expected in Tests A and B: Their main characteristics from a 'SOLO' perspective.

2.3.3 Application of the Tests

2.3.3.1 Subjects

Tests A and B were applied to three Spanish speakers with different levels of command of English, to explore the types of response that could be obtained in practice.

The following information about the subjects will be useful when considering their responses.

Subject A: Studied basic English in Spain. At the time of the application had been studying English for three months at a language school in Edinburgh (in a group of beginners).

Subject B: Studied basic English in Mexico. At the time of the application had studied English for several months in a language school in Edinburgh (at an intermediate level).

Subject C: Mexican post-graduate student of biology, with ample practice in reading texts of the field in English. Could therefore be considered advanced in terms of level of proficiency.

The three subjects were given Test A (text in English) first, and Test B (text in Spanish) afterwards. The transcripts of

their responses are given in Appendix C. An initial analysis of their responses was carried out, as described below.

2.3.3.2 Analysis of Subjects' Answers

SUBJECT: A.

TEST: A

This response is basically a multistructural one. Some of the basic information in the original passage is correctly reproduced in the summary. For example, some figures are quoted correctly, the symptoms and causes of some illnesses are correctly established, and some sources of information are also correctly identified. However, not even at this basic level is all the information included correct. For example, the questions with which the text begins were not "asked to the inhabitants of the city by a group of Mexican doctors." The high levels of noise in the city have not been called by the doctors "the metropolitan headache." The original passage does not state that "foreigners develop self-defence or adaptive mechanisms."

The doctors mentioned in the article do not state that "pollution affects the weather," nor that in turn "this may be the reason why temperatures are lower than in previous years."

Although an important relating concept is identified (possible risks of making existing information public), there is also a problem here, since the government is erroneously said to have "made up some statistics to avoid the danger."

SUBJECT: A

TEST: B

This response could be considered relational. All the basic

information included in the summary is correct. (There remains only one error; namely, that of not realising that the reporters ask the questions at the beginning, and that they are addressed to the readers). The relating concept that had not been recovered altogether correctly in Test A, is correct this time. Other integrating concepts present in the summary are the awareness of the limited action taken by the government, as well as of the lack of reliable official information.

SUBJECT: B

TEST: A

The only mistake in this summary is to have included "members of the National University and the National System for Family Integral Development" as well as the president of the country, in the list of the people interviewed. Apart from that, all the information it contains is correct.

Several relating concepts were recovered. Illnesses are correctly grouped by type of pollutant that causes them. It is pointed out that it is the members of the ecologist groups interviewed who provide the basic information, warn the public about the seriousness of the problem, and criticise the authorities for their lack of adequate research and action.

It is therefore a relational response.

SUBJECT: B

TEST: B

Again, the information included in this summary is correct, but this time there is no confusion at all about who was interviewed. The various relating concepts included in the previous summary were included in this one as well.

This time, besides a summary, Subject 'B' included comments on the material read. Actually, it is the comments that make this response qualify for the extended abstract category. A critical stance is taken. The article is qualified as "good" (though "short" and "with not much new information" for the subject). It is said to be a good article, because the sources of information referred to are reliable. Also, because it is seen as a potential awareness-raising tool, that may eventually lead to action. Finally, it is noted how politicians' and officials' interests do not coincide with the general public's welfare.

SUBJECT: C

TEST: A

All the information included in the summary is correct. Though it is short and concise, it is complete and structured in a clear way: Researchers of various institutions (correctly identified) are said to blame pollution in Mexico City for being the source of "new illnesses." Then, those illnesses are identified, along with those people who are more easily affected by them. The pollutants that cause them are then identified, as well as their sources. Finally, it is stated that although the government has some relevant information, it lacks statistical analyses, and that - what is more - that information is hidden from the public so as to avoid social problems.

This summary can be therefore considered relational. The use of quotation marks when referring to the "new illnesses," suggests that the subject does not share this view of the illnesses in question. However, since this view is not overtly stated, the summary can not be included in the 'extended abstract' category.

SUBJECT: C

TEST: B

The response of the subject in this test can be considered extended abstract. Like the previous summary, this one contains only correct, relevant information, and it is quite clearly written. The use of quotation marks makes it clear that the subject does not share the authors' conception of the illnesses as "new". Later, the subject distances himself from the text again to make it clear that it is the head of the MEM who thinks that the government is hiding information to avoid social unrest.

An overtly critical view of the article is given by the subject in the "comments" section of the response. The article is called "sensationalist". Data for which no statistical analyses exist is handled in the article. However, the subject recognises the importance of reporting the danger and its magnitude. Finally, the article is criticised for not including information about activities carried out by ecologist groups, nor about possible ways of alleviating the problem.

2.3.4 Comments

Although the answers of only three subjects were analysed, this exploratory study provided support for my intention of using the SOLO Taxonomy in the experimental phase of my research. On the one hand, the answers served the purpose of illustrating the type of response likely to be obtained. On the other, the analysis carried out was indicative of the suitability of the basic framework provided by the taxonomy for assessing such responses qualitatively.

Besides the comments to the subjects' responses included in the previous section, I would like to point out, regarding the "scores," that they all moved up one category in the second test. Subject "A" moved from the "multistructural" level to the "relational" one, and Subjects "B" and "C" from the "relational" to the "extended abstract." This, of course, may have been the result of several factors. For example, the facts that the second test was in the subjects' mother tongue; that they were already familiar with its content after having read the English text in Test "A", or an interaction of these.

This fact suggested the convenience of modifying the tests to be used in the experiment, and to control the order of their application too. As will be discussed in detail in the following chapter, it was decided that the texts to be read by the subjects should not be the same in the two tests. Also, it was decided that half the subjects in the experiment should take the test in English first and the one in Spanish second, while the other half should take the tests in the reverse order, thus controlling for the carry-over effect.

The convenience was also felt of using the complete SOLO Taxonomy including the transitional categories, rather than only the basic ones as in this preliminary study. It was anticipated that a wide range of responses was likely, and that the complete set of categories would facilitate and enrich their analysis.

The need became also apparent of giving clearer instructions to the subjects by asking them to produce their summary and their comments separately.

Finally, it was also clear that the marking scheme should be designed so as to reflect adequately this clearcut distinction between summaries and comments. For, as Hutchinson has noted too, the "extended abstract" level is not really appropriate in the case of summaries "where strict adherence to the information contained in the text is required." (Hutchinson, 1985, p.63).

On the other hand, this level may indeed be reached by the subjects in the "comments" section of the tests.

From the above discussion it will be clear that this exploratory study had important implications for the final design of my experiment, of which a complete account is given in the next chapter.

CHAPTER THREEThe Experiment

Chapter Three

The Experiment

3.1 Introduction

The experiment described in this chapter was designed taking into consideration insights gained from the exploratory study reported in Chapter 2. Two phases can be distinguished: the planning stage and the implementation stage.

The planning phase consisted of three stages. First, the final version of the research design was drawn up, defining design type, variables and hypotheses. These will be discussed in the next section.

The second stage was devoted to the preparation of the materials to be used. As far as the subjects' proficiency in the foreign language, it was decided to use two sources of information: the subjects' classification within the English Department of the Foreign Language Centre at the National University of Mexico, as well as that obtained from the application of selected sections of the English Language Battery Test (ELBA Test). Regarding the reading comprehension tests, as indicated in the previous chapter, it was necessary to develop a new battery of tests: two in English, and their corresponding translations into Spanish. Having designed the tests, a pilot study was carried out and the definitive marking scheme was formulated. Section 3.3 includes a detailed discussion of the materials and their preparation.

The third and final planning stage was that of selecting the subjects who would participate in the experiment. The actual selection took place in Mexico. Its procedure will be described in Section 3.4.

In the last section of the chapter (3.5), an account will be given of the procedure followed to apply and score the tests. The results are reported in the next chapter.

3.2 Research Design

Before giving a definition of the variables, hypotheses and type of design used, it may be convenient to start this section by including a brief description of the experiment.

As pointed out earlier, I am primarily concerned with analysing qualitatively the subjects' reading outcomes in English as a foreign language. However, they were also asked to carry out the same reading task in their mother tongue, Spanish, as a control. Since I am also interested in studying the way in which different levels of command of the foreign language may affect reading performance, subjects with differing degrees of proficiency in English were chosen for the experiment. Three levels were selected, and - for ease of identification - labelled "beginners," "intermediate" and "advanced."

To control for the possible carry-over effect, half the number of subjects took the English test first and the Spanish test second. The other half took the tests in the reverse order.

Figure 3.1 contains a schematic representation of the research design:

G1(B)	T1(E)	T2(S)
G2(B)	T1(S)	T2(E)
G3(I)	T1(E)	T2(S)
G4(I)	T1(S)	T2(E)
G5(A)	T1(E)	T2(S)
G6(A)	T1(S)	T2(E)

where:

G = Group	B = Beginners
	I = Intermediate
	A = Advanced
T = Test	E = English
	S = Spanish

Figure 3.1 Schematic Representation of Research Design.

With this initial information about the experiment, a definition of the variables studied is now in order.

3.2.1 Variables

The following variables can be identified in this study:

Independent Variable:

Hatch and Farhady point out that "it is the independent variable... which you believe will affect the other variable." (Hatch and Farhady, 1982,p.15).

In this study, therefore, the independent variable will be the "Language used in the test" (mother tongue, or foreign language), for it is the one believed to affect the subjects' reading outcomes.

Dependent Variable:

This is "the variable which you observe and measure to determine the effect of the independent variable." (Ibid.).

In this study the dependent variable will be the subjects' observed "Reading outcomes."

Moderator Variables:

This type of variable is "a special type of independent variable which you may select for study in order to investigate whether it modifies the relationship between the dependent and the major independent variables." (Ibid.).

In this study, there will be two moderator variables: "Level of proficiency in the foreign language" (Beginners, Intermediate, and Advanced), and "Order of administration of the tests" (FL first, L1 second; or, L1 first, FL second).

As mentioned before, the first one, "level of proficiency in the foreign language," was chosen because of the importance that studying the relationship proficiency-reading outcome has from a theoretical perspective (cf. Section 1.3.2.2).

On the other hand, the variable "order of administration of tests" was included due to more practical considerations regarding the research design. That is, as a result of the exploratory study, the convenience was felt to find a possible way to control the carry-over effect (cf. Section 2.3.4). There was no other interest in the actual order of administration of the tests than to eliminate its effect from the "language of test" and "error" variances. In other words, following Lindquist, this variable was introduced "only in order to make possible a valid test of the...["language

of test") differences, and to increase the precision of the test, and not because...[of interest] in the particular. . . ["orders of administration"] involved." (Lindquist, 1940,p.169). Moreover, by building this variable into the design, a way of studying "possible interactions [of this variable] with the experimental variable of major interest...is also provided for." (Kerlinger, 1967, p.235).

Having defined the variables involved, let us now turn our attention to the type of research design chosen.

3.2.2 Type of Design

The characteristics of the research design selected make it fall within the following categories:

Ex Post Facto Design:

Hatch and Farhady point out that ex post facto designs are used when "the researcher has no control over what has already happened to the subjects. The treatment, whatever it might be, has been given prior to the research project." (Op.cit., p.26). In the present study the subjects were not submitted to any specific language or reading training programmes in order to test their effects on the proficiency-reading outcome relationship being studied.

Higher Order Factorial Design:

Hatch and Farhady state that in factorial designs "there will be more than one independent variable (i.e., moderator variables) considered and the variables may have one or many levels." (Op.cit.,p.28).

More precisely, the type of design used in this study is a "Higher

Order Factorial Design," since it includes more than two independent variables: "language of test," "level of proficiency in the foreign language," and "order of administration of tests" (Cf. Spector, 1981, pp. 63-8).

Since "language of test" has two levels (L1 and FL); "level of proficiency," three (Beginners, Intermediate, and Advanced); and "order of administration," two (FL first, L1 second, and viceversa), the research design is a "2X3X2 higher order factorial design."

Repeated Measures Design:

As Norusis states, "the simplest repeated measures design is one in which two measurements are obtained for each subject - such as pre- and post-test scores." (Norusis, 1985, p.258).

Since in this study each subject appears under two of the experimental conditions (i.e., for each of the two levels of the "language of test" variable: L1 and FL), it may be said that it is also a repeated measures design.

Norusis points out the advantages of this type of design: "besides requiring fewer experimental units (in this study, human subjects), they provide a control on their differences." (Ibid.).

Experimental-Qualitative-Statistical Paradigm:

This is not a type of design, but I nevertheless want to include this way of classifying the present study here, because of the relationship between design type and "methodological paradigm." Grotjahn distinguishes eight such paradigms. The present study, it seems to me, falls within his fourth paradigm, whose characteristics are:

Paradigm 4: experimental-qualitative-statistical.

1. experimental or quasi-experimental design.
2. qualitative analysis.
3. statistical analysis.

(Grotjahn, 1987, p60).

As far as the first point is concerned, the type of design used in this study has already been mentioned above (ex post facto, higher order factorial, repeated measures design). As for the second and third points, it has already been stated that the data will be analysed qualitatively first (using as a basis the SOLO Taxonomy), and quantitatively afterwards (basically by means of the ANOVA statistical test).

Having examined the characteristics of the research design used in this study, the experimental hypotheses will now be considered.

3.2.3 Hypotheses

For reasons already stated (cf. Section 1.3.2.2) this study was basically set up to compare the reading performance of the same subjects in their mother tongue and the foreign language, but also to explore the way in which different levels of proficiency may affect the subjects' reading outcome in the foreign language.

The following questions have been asked in relation to such theoretical considerations:

- Are "good" readers in L1 also "good" readers in FL?
- Are "poor" readers in L1 also "poor" readers in FL?
- How may the reader's level of proficiency in the foreign language affect the quality of his reading performance in it?

These questions were formulated into Hypotheses 1 and 2 below. Hypothesis 3, on the other hand, stems from a more practical

consideration relative to the research design chosen (cf. Section 3.2.1).

- H_{0_1} : There is no significant difference in reading outcomes between L1 and FL.
- H_{a_1} : There is a difference in reading outcomes between L1 and FL, such that the subjects will do significantly better in L1.
- H_{0_2} : There is no significant difference in reading outcomes in the foreign language between the "Beginners," "Intermediate" and "Advanced" levels of proficiency.
- H_{a_2} : There is a significant difference in reading outcomes in the foreign language between the "Beginners", "Intermediate" and "Advanced" levels of proficiency, such that the higher the subjects' level of proficiency in English, the better they will perform in the English test.
- H_{0_3} : There is no significant difference in reading outcomes between the first or the second test taken.
- H_{a_3} : There is a difference in reading outcomes between the first or the second test taken, such that the subjects will perform significantly better in the second test they take.

The hypotheses are stated in the null form first as it is in this form that they will be subject to test. The alternative form follows, to make the direction of the expected difference clear. Following Robson's advice, the decision of using one-tailed

hypotheses was taken before doing the experiment (cf., Robson 1973, p.73). Also the decision of rejecting the null hypotheses at the conventional significance level of 5% was taken - in agreement with Rowntree - before collecting the data (cf., Rowntree, 1981, p.134).

The hypotheses stated above are related to the three main factors in the ANOVA test used to determine the statistical significance of the difference between the contrasting groups. It should be noted, however, that the ANOVA test also permits the exploration of possible interactions between such factors. These will be considered in Chapter 4.

In the next section, we shall now take into consideration those aspects of the investigation relative to the materials used.

3.3 Materials

3.3.1 FL Proficiency Test

To classify the subjects according to their proficiency level in English, two sources of information were available.

The first one was the subjects' classification within the English Department of the Foreign Language Centre at the National University of Mexico. First time students at the Centre are required to take a placement examination to allocate them in the appropriate classes. However, once students complete their first semester at the Centre, they simply move on to the next one the following semester, and so on. It can be noticed, then, that although

most of the students are properly placed - either by means of the initial classification exam, or by their marks at the end of each semester, - there is no single measure against which to compare all of the students in the department at any one time.

Consequently, it was decided to administer another test to all the participating subjects in the experiment - the same to all this time, - to ensure an additional, uniform source of information regarding their level of proficiency in English.

Selected sections of the English Language Battery Test (ELBA Test) were chosen for this purpose. Permission to use them was kindly granted by the Applied Linguistics Department of Edinburgh University.

The sections used were two: "Grammar" and "Vocabulary." Each of these sections is made up of 50 multiple choice items. They were applied and scored following the Instructions for Administration given in the Manual for the Test. (Ingram, 1975).

The way in which the subjects were classified using the above information about their level of proficiency in English will be discussed in Chapter 4.

3.3.2 Reading Comprehension Tests

Four reading comprehension tests were designed to be used in the experimental phase of this study: Tests 1 and 2 in English, and their corresponding translations into Spanish. The first aspect of their design to be discussed will be that of the selection of the English texts used.

3.3.2.1 Text Selection

After carrying out the exploratory study reported in Chapter 2, it was decided to use two texts for Tests 1 and 2 (cf. Section 2.3.4), to prevent the likely learning effect - and loss of interest - that reading the same text in English and Spanish might entail.

To tackle the problem of finding adequately matched texts, it was decided to use two consecutive sections of the same text. Thus, the texts in both tests would be by the same author, on the same topic, and from the very same source. This would allow a high degree of control regarding nature, content, style and degree of difficulty of the texts.

The texts used in Tests 1 and 2 (Appendices D and E respectively), then, are two excerpts from "The International Community and the Environment," by M. F. Strong (Strong, 1980, pp. 619-21). Apart from their being excerpts, the texts were not submitted to any other form of manipulation aimed at facilitating their use in an experimental situation.

The fact that the original was addressed to a public of non-specialists played an important role in the selection process, since the experimental subjects come from different academic backgrounds. Thus, it met the criterion of being accessible to readers without a technical knowledge of the subject.

On the other hand, the passages chosen were also "demanding enough and detailed enough to present a sufficient intellectual challenge." (Entwistle and Ramsden, 1983, p. 85). Thus, it was also expected that the passages would lend themselves to a wide range of response levels.

A related aspect of the passages also taken into consideration was that of topic. The ecological issues dealt with in the passages were expected to be of relevance to the subjects, and, accordingly, that their interest for reading them would be enhanced. The current importance of the topic was expected to be particularly useful in the "Comments" section of the tests, since it could prompt comments beyond the information presented in the passages, while their general nature made them open to comments from different perspectives (economic, medical or technological, for example).

Finally, from a formal point of view, it should be noted that both passages can be read as complete, independent units. As regards their length, (the text in Test 1 has 628 words; the one in Test 2, 662 words), the passages are neither too short to be summarized, nor too long that the subjects could not complete the required tasks in the time available (one hour).

Having chosen the English texts, it was necessary to translate them into Spanish to complete the battery of tests necessary for the experiment (cf.Section 3.2). The texts were translated by myself and another applied linguist, native speaker of Spanish. We first worked independently; then discussed our differences, and agreed on the final translation to be used in the Spanish versions of Tests 1 and 2 (Appendices F and G, respectively).

The next step was to conduct a readability study of the English and Spanish texts. This study will now be presented.

3.3.2.2 Readability Study

The purpose of the study reported in this section was to determine if the four texts to be used in the experiment (Tests 1 and 2, in English and Spanish) could be considered equivalent in terms of difficulty.

Although doubts about the use of readability formulae have been expressed (cf., for example, Taylor, 1953; and Irwin, 1979), there are also authors who support their use. So, for example, Nuttall points out that

Many measures of readability make use of counts of word length and sentence length; neither is a completely reliable indicator but the results give a rough guide which is useful if you have to assess the text yourself. More elaborate methods would take more time than most of us can afford.

(Nuttall, 1982, p. 26).

Furthermore, Davies - taking up a point made by Harrison (1979) - states that "the best measure of text difficulty is combined expert judgement and where that is not available, readability formulae." (Davies, 1984, p. 190).

With these ideas in mind, the Flesch readability formula (Flesch, 1948, and 1949) was chosen for the purpose of the present study. Harrison points out that Flesch found - after analysing many linguistic variables - "that two in particular seemed to correlate most highly with difficulty. These were the average number of syllables per word and sentence length." (Harrison, 1979, p. 77). Accordingly, his reading ease (RE) score (Fig. 3.2) is based on these two variables (ibid.; and Davies, 1984, p. 188).

$$RE = 206.835 - (0.846 \times NSYLL) - (1.015 \times W/S)$$

where: NSYLL = average number of syllables per 100 words, and
W/S = average number of words per sentence.

Figure 3.2 Flesch Readability Formula.

Flesch gives simple indications regarding the way in which the formula can be applied (1949,pp.247-51), and Harrison provides a clear example of the formula in use (1979,pp.106-7).

The results of applying the Flesch formula to extracts of around 300 words from the four texts used in this study are included in Table 3.1. The extracts in all cases were taken from the beginning of the texts up to the sentence where the 300th word occurred.

	Text 1	Text 2	Text 1	Text 2
	English	English	Spanish	Spanish
Number of Sentences	9	10	9	9
Number of Words	315	306	339	310
Number of Syllables	516	546	709	668
NSYLL	163	178	209	215
W/S	35	30	37	34
RE	33	25	-7	-9

Table 3.1 Flesch Reading Ease Scores of Texts 1 and 2, in English and Spanish. (The number of sentences, words and syllables are those used in computing the RE, not the total number in the complete texts).

It can be noticed that the number of sentences is the same in three of the texts, the second one in English having only one

sentence more than the others. The number of words is of course different, but there is a parallel variation: Text 1 has more words than Text 2 both in English and Spanish. The number of syllables is again different, though this time the main difference is not between Texts 1 and 2, but between the two languages: both Spanish texts have a larger number of syllables than their English counterparts. These differences in numbers of words and syllables of the texts are manifest in the NSYLL (average number of syllables per 100 words), and the W/S (average number of words per sentence) counts; and by definition determine the differences in their RE (reading ease score).

Since the formula was developed for texts written in English, it is hardly surprising that the main differences observed are again between the two languages. If an attempt were made to interpret the RE scores of the four texts in terms of the tables produced by Flesch (1949,p.177), the Spanish texts would indeed be much more difficult than the English ones. However, it is interesting to note that there is again a parallel variation regarding the RE scores; those for Text 2 (in English and Spanish) are lower than those for Text 1 (in English and Spanish). In spite of the interest that these observations may have in themselves, the fact that the texts were written in different languages made it necessary to submit the data used in the computation of the RE scores to some statistical tests to ensure as objective a judgement as possible regarding the observed differences.

The first test used was the Wilcoxon test for small samples. This test is the non-parametric counterpart to the correlated

t-test, and is used to test whether or not the populations are identical (cf. Robson, 1973, pp.106-14). The computation of this test was based on all the information available for each of the texts regarding their RE scores. That is, each pair of texts was compared taking into consideration their number of sentences, number of words, number of syllables, NSYLL, W/S, and RE.

As can be seen in Table 3.2, the comparisons were made between Texts 1 and 2 in English, Texts 1 and 2 in Spanish, Text 1 in English and Spanish, and Text 2 in English and Spanish. No significant differences were found in any of these cases.

TEXTS	OBSERVED T	VALUE OF T (WHEN N = 6)	COMMENTS
Text 1 English- Text 2 English	9	1	As the observed values of T are greater than the corresponding table values (Robson, 1977, p.153), there is no significant difference between the two conditions in each of the comparisons being made.
Text 1 Spanish- Text 2 Spanish	5	1	
Text 1 English- Text 1 Spanish	4	1	
Text 2 English- Text 2 Spanish	4	1	

Table 3.2 Results of the Wilcoxon Test Applied to Different Pairs of the Texts in the Study.

A second test was then used, this time to compare all four texts on a single variable at a time. The χ^2 test was chosen for this purpose. Tables 3.3, 3.4 and 3.5 include the data used in the computation of χ^2 for each of the following variables: NSYLL (average number of syllables per 100 words), W/S (average number of words per sentence), and RE (reading ease score).

	Text 1	Text 2
English	163	178
Spanish	209	215

Table 3.3 NSYLL of Texts 1 and 2, in English and Spanish.

	Text 1	Text 2
English	35	30
Spanish	37	34

Table 3.4 W/S of Texts 1 and 2, in English and Spanish.

	Text 1	Text 2
English	33	25
Spanish	-7	-9

Table 3.5 RE of Texts 1 and 2, in English and Spanish.

Table 3.6 summarizes the results of the χ^2 test applied to the above data. Since in none of the cases does the χ^2 obtained exceed the table value (cf. Robson, 1977,p.151), there is no significant evidence for an association between the categories ("language" and "text") in terms of their NSYLL, W/S and RE values at the 5% significance level.

VARIABLE	OBSERVED χ^2	VALUE OF χ^2 (WITH 1 d.f.)
NSYLL	0.1168	3.841
W/S	0.0004	3.841
RE	2.0678	3.841

Table 3.6 Results of the χ^2 Test Applied to the NSYLL, W/S and RE Values of the Texts in the Study.

Thus, since the results of the Wilcoxon and χ^2 tests did not establish a significant difference between the four texts regarding their degree of difficulty, it was decided that they could be used as intended.

Having taken this decision, it was now necessary to analyse the texts from a SOLO perspective as a first step towards the preparation of the marking scheme. An account of this analysis follows now.

3.3.2.3 Analysis of Texts 1 and 2 from a SOLO Perspective

To follow Biggs and Collis's recommendation to analyse in advance the task the subjects are required to carry out (Biggs and Collis, 1982, pp.165-72), let us first of all remember that the subjects' reading tasks in this study were to state in their own words the topic of the passage, and to make comments on the materials read, separately. (See "Instructions", Appendix D). This analysis will serve the purpose of identifying the type of information that can count as relevant for each of the levels in the taxonomy. (See also what has already been discussed in this respect in Section 2.3.2).

As already mentioned, these tasks were expected to prompt a wide range of response levels, this variety of outcomes being quite convenient for the qualitative analysis intended.

These tasks are also realistic in the sense that the subject population can be reasonably expected to summarize and make comments on this type of text in the course of their studies.

Fawcett (1979, p.214) has also pointed out that asking subjects to summarize a text is a more realistic task than to ask them to answer a set of predetermined comprehension questions, in the sense that in the first case the reader reads to inform himself, while comprehension questions distort the nature of the reading. Lunzer, White and Dolan (1970, p.69) also note regarding the usefulness of this task that Thomas and Augstein (1972) "found that reading a passage in the expectation of a request to write a summary produced better results than reading the same passage in the expectation of a comprehension test."

These tasks also come closer to the type of question that specialists appear to prefer, as opposed to those asked by the non-specialist in the subject (e.g. language teachers). As Zuck and Zuck found:

The answers to the questions posed by non-specialists tended to be more localized in the text, often based on the information in a single sentence or single paragraph. In contrast, the specialists tended to ask questions based on an interpretation of larger units of the texts, or even the text as a whole.

(Zuck and Zuck, 1984, pp.134-5).

On the other hand, the convenience of asking the subjects to produce their summaries and comments separately was clear after conducting the exploratory study reported in Chapter 2, and is closely related to the nature of the instrument used to analyse them. Since in a summary the subjects must confine themselves to the data in the original, whereas in the "comments" section they are expected to go outside the text, the first levels in the SOLO Taxonomy are adequate for analysing the summaries, and the extended abstract category can only be found in the "comments" section of the tests. This fact has also been noted by Hutchinson (1985, p.63).

To make it clear that they were required to produce summaries and comments separately, the subjects were given their instructions in Spanish; and given their different levels of proficiency in English, they were asked to answer in Spanish as well. (See Appendix D).

Since in tasks like these the subjects' responses are directly related to the specific materials used, Texts 1 and 2 were analysed from a SOLO perspective. The analysis will provide - as previously noted - a list of "components or dimensions comprising the task

that enable the distinctions between the various SOLO levels to be made." (Biggs and Collis, 1982, p. 165). This procedure had already proved to be useful in the exploratory study reported in Chapter 2. In fact, the analysis of the materials used in that study (Section 2.3.2) provided useful guidance to analyse Texts 1 and 2 as reported below. Since the Spanish texts are translations of the English ones, the analysis that follows is valid for the texts in either language. (The texts are included in Appendices D-G).

Table 3.7 includes the analysis of Text 1. The first column gives the labels of the different levels. The second one gives information about the way in which such levels can be established according to the authors of the taxonomy. In the third one, a specification is made of the characteristics of the responses that would fall into each level on the basis of the information presented in the text.

LEVEL OF RESPONSE	MAIN CHARACTERISTICS OF RESPONSES DEPENDING ON LEVEL	MAIN CHARACTERISTICS OF RESPONSES THAT MAY BE EXPECTED IN TEST 1
P r e s t r u c t u r a l	The basic information in the passage has not been recovered correctly.	Responses at this level are very unlikely. However, responses in this category would be those showing the subject's lack of understanding of: - what the original passage is about (e.g., mistaken information included in the summary), and, or: - what the task requirements are - writing a summary <u>and</u> their own comment- (e.g., confusion between facts stated in the text and the subject's own opinion - or extra information on the subject).
U n i s t r u c t u r a l	Any of the basic facts, details and concepts that are given in the original passage would constitute a unistuctural response (if addressed the requirement appropriately).	Responses at this level would be those that would deal with just one aspect of the original passage. For example, responses that recover correctly what the original text states regarding population growth in the developing countries, but do not go beyond that, would fall within this category.
M u l t i s t r u c t u r a l	Several of the basic facts, details and concepts that are given in the original passage could make up a multi-structural response (if addressed the requirement appropriately).	Responses in this category would fail to establish clearly the way in which the data in the original are interrelated. These responses though conveying correct information, fail to give a coherent view of the whole original passage. An example could be a summary including information about population growth in developing countries, and mentioning the need to control it, but without relating this aspect of the original passage to the point made by the author about the need to control man's growing demands for material goods in order to achieve a balance between man and the natural environment.

R e l a t i o n a l	An integrating concept in the original passage is recovered.	Responses in this category should establish clearly the way in which the data in the original are interrelated. Relating concept: "Both population growth and the growing demand for material goods must be controlled in order to achieve a balance between Man and the natural environment."
E A x b t s e t n r d a e c d t	Data, principles, concepts, hypotheses, that are not given in the original passage, but which are relevant are included in the response.	The form an extended abstract response will take is by definition unpredictable, and can only be expected in the "Comments" sections of the tests, since the subject goes beyond the text, reflects upon the material read, takes a critical stance and makes a personal contribution.

Table 3.7 Main characteristics of the responses that may be expected in Test 1 (English and Spanish), from a SOLO perspective.

Table 3.8 is the corresponding table for Text 2. To avoid repetitions, the basic information about the different categories included in Table 3.7 has been omitted. Therefore, only some examples specifically related to Text 2 are included this time.

LEVEL OF RESPONSE	RESPONSES THAT MAY BE EXPECTED IN TEST 2
Prestructural	Same as for Test 1
Unistructural	For example, responses that do not go beyond a mere enumeration of the instances mentioned in the original of ecological damage to the environment.
Multistructural	An example could be a response giving examples of ecological damage in support of the "no-growth" approach advocated by some environmentalists, without relating such information to the major argument of the text (cf. the "relating concept" in the next category).
Relational	Relating concept: Against the present approach to growth (based on the degradation and using up of the natural environmental capital), and a "no-growth" approach is proposed (based on the removal of the artificial and self-defeating conflict between economic and environmental factors).
Extended Abstract	Same as for Test 1.

Table 3.8 Responses that may be expected in Test 2 (English and Spanish), from a SOLO perspective.

As already mentioned, the above analyses were made to provide the basic framework to analyse the subjects' reading outcomes.

In addition, Biggs and Collis (1982,p.170) point out the advisability for teachers (or researchers) of completing the intended task for themselves before asking the subjects to do so. This way, one can gain a better understanding of the relationship between the task requirements and the expected outcomes. With this purpose in mind, this experimenter and two other specialists summarized both texts. The results of that phase of the study are given in the next section.

3.3.2.4 Specialists' Summaries of Texts 1 and 2

Two reading specialists - this experimenter and another applied linguist - (specialists A and B, respectively), and a content specialist - a biologist with experience in ecophysiology - (specialist C), summarized Texts 1 and 2. The specialists are native speakers of Spanish, and their command of English allowed them to be Ph D students at Edinburgh University in their respective areas of specialisation.

The purpose of this phase of the study was to find out if there was agreement among these specialists as to what this "particular task involves and what contents are crucial and what are not," as Biggs and Collis (1982,p.170) advise. This way, too, the scheme just developed for analysing the subjects' reading outcomes (cf. Section 3.3.2.3) could be tried out for the first time.

The texts were read in English, and the summaries written in Spanish (Appendix H). They were marked by two judges independently. Since the judges were also the reading specialists, the scripts that they had themselves written were marked only by one judge, i.e., not by themselves. The ones they did mark had been masked to avoid identification.

Table 3.9 includes the results. As can be seen, there was agreement between the two judges. As was to be expected, the three specialists reached the highest possible level in the taxonomy for this task (relational), supporting the view that they could agree on what this particular task involved.

		JUDGE A	JUDGE B
TEXT 1	Specialist A Specialist B Specialist C	-- Relational Relational	Relational -- Relational
TEXT 2	Specialist A Specialist B Specialist C	-- Relational Relational	Relational -- Relational

Table 3.9 Specialists' Scores in Summaries of Texts 1 and 2.

Having found no particular problems in this phase of the study, it was now necessary to try the materials out in a pilot study. This study will be reported in the section that follows.

3.3.2.5 Pilot Study

To pilot the materials, a small scale study was conducted with the participation of eight subjects as similar as it was possible to find in Edinburgh to those who would be the actual experimental subjects. Most of them were Mexican, though Subjects 4 and 8 were Panamanian and Spanish, respectively. They all had different academic backgrounds. Six of them were post graduate students at Edinburgh University. Subject 5 had just completed her BA, and had therefore satisfied the language requirements established by her university in Mexico, i.e., a specialised reading comprehension test in English. Subject 7 was about to complete his BA, and was studying an advanced English course in a language school in Edinburgh at the time of the pilot study.

The objectives of the pilot study were:

1. To check if the instructions were clear. (Appendix,I).
2. To check if the subjects could complete the tests in the time available.
3. To check if the type of expected responses were actually observed.
4. To get some practice in the application of the proposed marking scheme.
5. To get some initial information regarding the interjudge reliability of the scoring procedure.

In order to approximate as much as possible to the research design for the experiment (cf. Section 3.2.2), and to control for order of administration, some subjects did Test 1 in English and Test 2 in Spanish, while others took the tests in the reverse order (Table 3.10). No attempt was made later on, however, to test for the statistical significance of the differences observed between these two groups of subjects since they were too small, and because that was not one of the objectives of the pilot study.

SUBJECT	TEST 1	TEST 2
1	English*	Spanish
2	English	Spanish*
3	English	Spanish
4	English	Spanish
5	Spanish	English
6	Spanish	English
7	Spanish*	English
8	Spanish	English*

Table 3.10 Tests Taken by Subjects in Pilot Study. Those with an asterisk are included in Appendix J.

The subjects' scripts were marked by this experimenter and the reading specialist who had already marked the other specialists' summaries (cf. Section 3.3.2.4). The scripts were masked to avoid potential marker bias. A representative sample of the scripts can be found in Appendix J (those with an asterisk in Table 3.10).

Table 3.11 shows the marks that were allotted to the different levels of the SOLO Taxonomy used to assess the subjects' reading outcomes. Following Biggs and Collis (1982, pp. 203-4), transitional responses are quantified as half points. The definition of the categories used has already been given in previous sections: Tables 3.7 and 3.8 (Section 3.3.2.3) include the definition of the main categories for Tests 1 and 2 respectively, and Figure 2.2 (Section 2.2) includes that of transitional responses.

MARK	SOLO LEVEL
1	Prestructural
1.5	Transitional
2	Unistructural
2.5	Transitional
3	Multistructural
3.5	Transitional
4	Relational
4.5	Transitional
5	Extended Abstract

Table 3.11 Marks Allotted to SOLO Levels.

The subjects' summaries and comments were scored separately, but since the comments are not in fact independent from the summaries, only the highest of the two marks obtained by each subject in each test was used in the computations to estimate interjudge reliability. Accordingly, Tables 3.12 and 3.13 show the highest mark obtained by each subject in Tests 1 and 2 respectively, according to Judges A and B. They also specify the degree of difference between the two judges for each of the cases.

SUBJECT	JUDGE A	JUDGE B	DIFFERENCE BETWEEN A AND B
1	4.5	5	.5
2	4.5	5	.5
3	3.5	3.5	0
4	3	3	0
5	4	3.5	.5
6	3.5	3.5	0
7	3	3	0
8	3.5	4	.5

Table 3.12 Subjects' Marks in Test 1:
Differences Between Two
Judges. (Pilot Study).

SUBJECT	JUDGE A	JUDGE B	DIFFERENCE BETWEEN A AND B
1	5	4	1
2	4.5	4.5	0
3	3.5	4	.5
4	2.5	3	.5
5	4	3.5	.5
6	5	5	0
7	4.5	4.5	0
8	5	4	1

Table 3.13 Subjects' Marks in Test 2:
Differences Between Two
Judges (Pilot Study).

The degree of agreement between the judges is specified in Table 3.14. In seven cases there was complete agreement, in seven more there was a difference of only half level, and in only two cases the difference was of one level.

	AGREE	HALF LEVEL DIFFERENCE	ONE LEVEL DIFFERENCE
N	7	7	2
PERCENTAGE	43.75	43.75	12.25

Table 3.14 Degree of Agreement Between Two
Judges on Tests 1 and 2.
(Pilot Study).

Finally, the Pearson product-moment correlation coefficient was calculated: first for each of the tests, and then for both of them together (Table 3.15). The statistical significance of these coefficients was determined by checking the correlation table in Hatch and Farhady (1982,p.277). The three observed values were checked against those provided for the corresponding degrees of freedom for two-tailed tests. They were all significant; more so, when the results of the two tests were considered together due to the corresponding increase of degrees of freedom.

TEST	d.f.	r	p
TEST 1	6	.9126	.01
TEST 2	6	.7435	.05
TESTS 1 AND 2	14	.7806	.001

Table 3.15 Correlation Between Two Judges
on Test 1, Test 2 and Tests
1 and 2. (Pilot Study).

Having completed the study, it was possible to assert that its objectives had indeed been achieved. Regarding the first objective mentioned above, it became manifest that the instructions had to be made more explicit (the instructions "before" and "after" are included in Appendices I and D, respectively). Since in four occasions the subjects did not include their comments, and in two cases they were intermingled in the summaries, it was necessary to re-write the instructions making it clear that both summary and comments were required, and that they were to be produced separately.

It was also possible to observe that the time available for the tests was enough.

As far as the third objective is concerned, it was also possible to notice that there was agreement between expected and observed outcomes. In other words, it seemed that the expected levels were realistic in terms of the subjects' capabilities. It was also reasonable to suppose that this could again be the case with the actual experimental subjects.

It was also possible to get the expected practice in the application of the taxonomy to the analysis of the subjects' reading outcomes. In fact, it became apparent to both judges that its application was not really difficult.

Finally, and perhaps more importantly, it was possible to get encouraging results as far as interjudge reliability of the marking scheme. It was decided, therefore, to use it without modifications to mark the experimental subjects' responses.

In the next section, information will be given about the subjects who participated in the experimental phase of this investigation.

3.4 Experimental Subjects

As already mentioned, the subjects who participated in the experiment were studying English as a foreign language at the Foreign Language Centre (CELE) of the National University of Mexico (UNAM). They did not share the same academic background, for the approximately 2000 students who study English at the Centre each semester come from all fields of study offered at the university. Each

semester they are admitted to the corresponding one of the nine levels that make up the English language programme either on the basis of the results of a placement examination, or - in most cases - by moving up the ladder from one level to the next. The semester in which the experiment was carried out - the second semester of the academic year 1986/87 - there were 2059 students in the English Department of the Centre (cf. Table 1.1, Section 1.2).

The subjects were chosen at random, using a multistage sampling procedure. This procedure, "as the term suggests, consists of sampling by successive stages." (Butler, 1985,p.7). In it, "the population is divided into...large sampling units and a random sample of smaller units is drawn." (Reichmann, 1964,p.252).

In the first stage of the sampling procedure, then, the population was divided into three main sampling units: "Beginners" (levels 1-3), "Intermediate" (levels 4-6), and "Advanced" (levels 7-9). From each of these units, three classes were selected at random. This yielded a total of nine classes, with 183 students (Table 3.16, Columns 1 and 2). This initial number of students was considered appropriate to ensure that a sufficiently large number of subjects could complete the battery of tests required for the experiment. However, after the administration of the tests had been completed, a listwise missing-value treatment was used. This treatment means that "a case is eliminated if it has a missing value on any variable in the list" (Norusis, 1985,p.36). This way 72 subjects were eliminated (Table 3.16, Column 3). Of the 111 who completed the battery of tests (Column 4), 58 took Test

1 in English and Test 2 in Spanish, while 53 took the tests in the reverse order (Columns 5 and 6). The subjects were assigned at random to these conditions when the English and Spanish versions of the tests were distributed in each class.

Class	Number of Subjects (Initial)	Missing Values	Subjects who took the three tests	Test 1 English/ Test 2 Spanish	Test 1 Spanish/ Test 2 English
Beginners 1	22	12	10	6	4
Beginners 2	14	2	12	6	6
Beginners 3	27	10	17	9	8
(Partial Totals)	(63)	(24)	(39)	(21)	(18)
Intermediate 1	20	13	7	4	3
Intermediate 2	19	3	16	8	8
Intermediate 3	29	12	17	7	10
(Partial Totals)	(68)	(28)	(40)	(19)	(21)
Advanced 1	19	6	13	9	4
Advanced 2	18	10	8	5	3
Advanced 3	15	4	11	4	7
(Partial Totals)	(52)	(20)	(32)	(18)	(14)
TOTALS	183	72	111	58	53

Table 3.16 Number of Subjects Participating in the Experiment, by Level of Proficiency (CELE Classification), and Order of Administration of the Tests.

In the second stage, the results of the ELBA Test were also taken into consideration (cf. Section 4.2). With this information about the subjects' proficiency in English at hand, some modifications were made to the "Beginners," "Intermediate" and "Advanced" groups that had been formed according to the CELE classification. The 111 participating subjects were divided into three groups of equal number (37) at the 33rd and 66th percentiles. (Cf. Martuza, 1977, p.33). Thus, the subjects were divided into "Beginners" (low 33.3%), "Intermediate" (middle 33.3%), and "Advanced" (top 33.3%), with regard to their performance in the ELBA Test (Table 3.17).

LEVEL	TEST 1 ENGLISH/ TEST 2 SPANISH	TEST 1 SPANISH/ TEST 2 ENGLISH	TOTALS	
			N	%
Beginners	21	16	37	33.3%
Intermediate	19	18	37	33.3%
Advanced	18	19	37	33.3%
TOTALS	58	53	111	100%

Table 3.17 Number of Subjects who Completed the Battery of Tests in the Experiment, by Level of Proficiency (ELBA classification), and Order of Administration of the Tests.

To keep the number of subjects equal within each of the experimental conditions, to save computational effort (since I did not intend to use a computer when carrying out the statistical analysis of the data), 15 subjects were finally selected at random from the available subjects in each condition (Table 3.18).

LEVEL	TEST 1 ENGLISH/ TEST 2 SPANISH	TEST 1 SPANISH/ TEST 2 ENGLISH	TOTALS	
			N	%
Beginners	15	15	30	33.3%
Intermediate	15	15	30	33.3%
Advanced	15	15	30	33.3%
TOTALS	45	45	90	100%

Table 3.18 Experimental Subjects, by Level of Proficiency and Order of Administration of Tests.

It was these subjects whose reading outcomes were submitted to the qualitative and quantitative analyses reported in Chapter 4.

Before presenting the results of such analyses, an account of the procedures followed in the application and scoring of the tests will be given in the section that follows.

3.5 Procedure

Two procedural aspects of the investigation will be taken into consideration next. The procedure followed in the application of the tests will be described first, and then, that followed to score them.

3.5.1 Application of Tests

Once the classes that would participate in the experiment were selected in agreement with the Head of the English Department (cf. Section 3.4), the teachers were informed about the nature and purpose of the study, and they all kindly agreed to lend their classes. Having obtained the teachers' permission, the students were approached. In all cases, this experimenter explained to them the purpose and nature of the study. They were also informed about the number and type of tests they would be required to take. Their cooperation was asked for for three sessions, one each of the following weeks. The importance that their attendance to the three consecutive sessions had for the experiment was made clear to them, as well as the fact that the results would only be used for the purposes of the experiment and would not count for their course marks.

During the following three weeks the actual application of the tests took place. In order to ensure that the conditions of application of the tests were the same for all participating subjects, they were always applied by this experimenter.

Test 1 was applied during the first week. The English and Spanish versions of the tests were randomly applied to half the number of subjects in each class. Test 2 was applied the second week, the English and Spanish versions being given to the corresponding subjects. During the third week the ELBA Test was applied to all of the subjects.

The same procedure was followed for the application of Test 1 and Test 2:

1. The tests were distributed.
2. The students were asked to make sure they provided the necessary information to identify them in the corresponding answer sheets.
3. The tasks were made clear to them: a summary (not a translation) of the text, and their comments were required.
4. It was also made clear that they had separate answer sheets for their summaries and comments.
5. They were asked not to use their dictionaries.
6. It was pointed out that they had a time limit of 50 minutes.
7. If they had any doubts, they were solved at this stage. Otherwise, they were asked to start working.
8. Ten minutes before the time limit, they were advised about the time so they could organise their work during the time available.
9. Those students who completed the test early, handed in their papers when they finished. Otherwise, I collected them at the end of the 50 minutes allowed (this, by the way, was the case of a minority of students in each class).

The procedure followed for the application of the ELBA Test was that specified in the Manual for the Test (Ingram, 1975).

One week afterwards, the subjects received feedback about their

performance on the ELBA Test (they knew it would not be possible for me to give them any feedback about the reading comprehension tests), and were duly thanked for their cooperation in the experiment.

3.5.2 Scoring of Tests

The scoring of the ELBA Test was carried out by using the corresponding key.

As for the reading comprehension tests, they were scored as follows. The subjects' scripts were masked, and marked by two judges (this experimenter and the other judge who had marked the tests in the pilot study), working independently. As already mentioned (cf. Section 3.3.2.5), after conducting the pilot study, it was decided to use the same marking scheme to mark the experimental subjects' responses. The definition of the categories used as applied to these particular tests can be found in Section 3.3.2.3 (Tables 3.7 and 3.8). The definition of the transitional levels was that given by Biggs and Collis (1982,p.36) (cf. Figure 2.2, Section 2.2). The marks allotted to the SOLO levels were the same as in the pilot study (cf. Table 3.11, Section 3.3.2.5).

As in the pilot study too, the subjects' summaries and comments were scored separately, but since the comments are not independent from the summaries, only the highest of the two marks assigned by each judge to the subjects in each test was used in the statistical analyses - as reported in Chapter 4.

Before carrying out the statistical analyses of the data, however,

those cases where there were any doubts of major differences between the two judges were discussed. This discussion proved very useful not so much in terms of the modifications made to some of the marks, but rather in terms of a better understanding of the subjects' reading outcomes and the instrument used to mark them.

After the discussion already mentioned, the definitive marks were recorded, thus completing the experimental phase of the investigation. In the chapter that follows, the corresponding results will be presented.

CHAPTER FOURResults

Chapter Four Results

4.1 Introduction

The results of the experiment will be presented in three separate sections. In the first one, consideration will be given to the subjects' performance in the English proficiency test (ELBA scores). The relationship between these scores and the subjects' classification within the English Department (CELE scores) will be examined, as well as the procedure followed to form the contrasting groups required for the experiment according to the subjects' level of proficiency in the foreign language ("Beginners," "Intermediate" and "Advanced").

The second section will be devoted to the results of the qualitative analysis of the subjects' reading outcomes. First, the SOLO scores assigned by two independent judges to the subjects' responses in Tests 1 and 2 (in English and Spanish, respectively) will be presented. Then, the degree of agreement between the two judges will be estimated, thus examining the central issue of interjudge reliability.

The final section of the chapter will deal with the results of the main quantitative analysis of the data, i.e., the results of an ANOVA test used to determine the statistical significance of the differences observed amongst the contrasting groups. As already mentioned, the research design used in this study is a three-way design. Correspondingly, a three-way ANOVA was applied to the data. The independent variables involved were "Language of Test," "Level of Proficiency" and "Order of Admin-

istration," and the dependent variable was "Reading Outcome." The results of this test will be presented in two parts. First the results of a complete ANOVA will be considered, and then, those of two breakdown ANOVA's used to further analyse the interaction "Language of Test" x "Level of Proficiency," which was the only one of the interactions in the complete ANOVA that proved significant.

4.2 Proficiency Levels

It has already been pointed out (Section 3.4) that out of 183 subjects who began participating in the experiment, 111 took the complete battery of tests (Reading Comprehension in Spanish, Reading Comprehension in English, and Proficiency in English), and that they were divided into three groups according to their CELE classification ("Beginners," "Intermediate" and "Advanced"). The scores obtained by the subjects in these groups in the ELBA Test are shown in Table 4.1.

	NUMBER OF SUBJECTS			
	(CELE CLASSIFICATION)			
ELBA SCORES	"ADVANCED"	"INTERMEDIATE"	BEGINNERS"	TOTALS
69	1			1
56	1			1
55	1			1
52	1			1
45	1			1
42		1		1
41	2			2
40	1			1
37		2		2
36	2			2
34	1	1	1	3
33	1			1
32		2		2
31	3	3		6
30	3			3
29	4	3		7
28		2		2
27	1	1		2
26		1		1
25		1		1
24	3	3	1	7
23	1	3	2	6
22	1		2	3
21	1	1	3	5
20	1	1	1	3
19		3	4	7
18		2	1	3
17	1	2	3	6
16	1	3	2	6
15			2	2
14		3	5	8
13		1	1	2
12			1	1
11		1	4	5
10			1	1
8			1	1
7			2	2
6			1	1
4			1	1
TOTALS	32	40	39	111

Table 4.1 ELBA Scores, According to Subjects' CELE Classification.

It has also been mentioned before (Section 3.3.1) that the sections of the test applied were two, "Grammar" and "Vocabulary," each with 50 multiple choice items. The maximum possible score was therefore 100. Since I did not use the whole test, no appeal can be made to the "official" statistics of the test to analyse the results of its application in this study. It should also be remembered that the test was basically used to be able to count on a single measuring device to compare all the participating subjects' proficiency in English at the time of the experiment (cf. Section 3.3.1). Thus, the scores shown in Table 4.1 must be taken at face value (e.g., one subject got a score of 4 - out of a maximum possible of 100 points in the two sections of "Grammar" and "Vocabulary" - another got a score of 6, two more got 7 points, and so on), and considered only in relation to the sampled population.

As indicated in Section 3.4, the "Advanced," "Intermediate" and "Beginners" groups formed according to the subjects' CELE classification were rearranged on the basis of their ELBA scores. Instead of having groups of unequal size (32, 40 and 39 subjects, respectively), with no clear-cut distinction regarding ELBA scores (cf. Table 4.1), it was decided to divide the 111 subjects into three groups of equal size ($n=37$), at the 33rd and 66th percentiles of the distribution. The scores obtained by the subjects in the groups thus formed in the ELBA Test are shown in Table 4.2. It must be pointed out that of the three subjects who scored 18 points, two were sent to the "Intermediate" group and one to the "Beginners" group. This was done at random.

ELBA CLASSIFICATION		SUBJECTS	
GROUP	SCORES	NUMBER	%
"Advanced"	69 - 28	37	33.3
"Intermediate"	27 - 18	37	33.3
"Beginners"	18 - 4	37	33.3
TOTAL		111	100%

Table 4.2 Number and Percentage of Subjects in the "Advanced," "Intermediate" and "Beginners" Groups, According to their ELBA Scores.

In order to test the degree of agreement between the two sources of information available regarding the subjects' level of proficiency in English (their CELE classification and ELBA scores), a χ^2 test was applied to the data (Table 4.3).

		CELE CLASSIFICATION			TOTALS
		"ADVANCED"	"INTERMEDIATE"	"BEGINNERS"	
E S L O B R A E S	69-28	22	14	1	37
	27-18	8	16	13	37
	18- 4	2	10	25	37
TOTALS		32	40	39	111

Table 4.3 Number of Subjects, by CELE Classification and ELBA Scores.

Following Robson's procedure for estimating it (Robson, 1973, pp. 94-100), the value of χ^2 for the above data was obtained: 43.31. Since this value is more than the corresponding value (i.e., when there are 4 degrees of freedom, as in this case) in the Table of χ^2 provided by Lindquist (1953, p.29) at the .001 probability level (i.e., $\chi^2 = 18.465$), there is significant evidence for an association between the variables involved (i.e., CELE classification and ELBA scores).

In other words, the discrepancies observed between the "Advanced," "Intermediate" and "Beginners" groups formed according to the subjects' CELE calssification, or their ELBA scores, are not statistically significant. Either system of classification could have been used indistinctly. The one finally chosen (ELBA scores), though, made it easier to form groups of equal size, on the basis

of a single measuring instrument.

Having decided to use the groups formed according to the subjects' ELBA scores, it was then necessary to check if their differences were significant enough to use them experimentally. In other words, it was necessary to test the null hypothesis of no significant difference among the "Beginners," "Intermediate" and "Advanced" groups regarding ELBA scores. A one-way ANOVA test was in order, since there were three groups to be compared, on one dimension only.

Using information from Tables 4.1 and 4.2, the mean (\bar{X}), standard deviation (SD), variance (Var) and sum of squared deviations ($\sum (X - \bar{X})^2$), were calculated for each of the groups being contrasted. The resulting figures, shown in Table 4.4, provided the data for the one-way ANOVA in question.

	Beginners	Intermediate	Advanced
n	37	37	37
\bar{x}	13.32	21.91	35.70
SD	3.4645	2.4426	9.1085
Var	12.0029	5.9663	82.9656
$\sum (x - \bar{x})^2$	444.10	220.76	3096.73

Table 4.4 Data for One-way ANOVA, ELBA Scores.

The procedure followed to run the ANOVA test on the data was that found in Reichmann (1964, pp.330-4). The results are summarized in Table 4.5.

Source of Variation	Sum of Squares	d.f.	Variance Estimate	<u>F</u>
Between Samples	4716.02	2	2358.01	68.20***
Within Samples	3734.59	108	34.57	
Totals	8450.61	110		

Table 4.5 ANOVA for Subjects' ELBA Scores.

Since the observed value of \underline{F} is larger than the critical value of \underline{F} for 2/108 d.f. for the .001 level of significance (cf. Table 3, Lindquist, 1953, p.44), we can reject the null hypothesis. The difference between the "Beginners," "Intermediate" and "Advanced" groups is highly significant.

Once it was established that the difference in proficiency in English among these groups was statistically significant, the experimental groups were formed by selecting 30 subjects at random from those available in each group (cf. Section 3.4), making sure that half the number of subjects in each group had taken the reading comprehension test in English first and in Spanish second, while the remaining half had taken the tests in the reverse order (cf. Table 3.18). The experimental subjects' ELBA scores and corresponding classification can be found in Appendix K.

4.3 SOLO Levels

4.3.1 SOLO Scores

Following the procedure indicated in Section 3.5.2, the subjects' scripts were marked by two independent judges. Those cases in which there was a difference of more than one level between the marks awarded by the judges were discussed, agreeing to modify them in some cases (cf. Section 4.3.2). The subjects' definitive scores are given in Appendix L. The discussion of the subjects' responses in terms of content will be included in Chapter 5.

Figures 4.1 through 4.12 show the graphic distribution of the SOLO scores obtained by the subjects in each of the experimental conditions, and Table 4.6 summarizes the results. The scores being taken into consideration are the highest ones obtained by each subject in each test, regardless of which of the two judges was the one who awarded it (Appendix L, Column VI). These scores were chosen - rather than the average of the two judges marks - for the purpose of reporting the results of the qualitative analysis of the subjects' reading outcomes since they can be readily interpreted in SOLO terms (cf. Table 3.11, Section 3.3.2.5, and Figure 4.1). In contrast, averages sometimes include quarters of points, which are not very meaningful in SOLO terms. On the other hand, averages were chosen for the main quantitative analysis of the data - the ANOVA test reported in Section 4.4 - because they capture more information about the two judges' assessment.

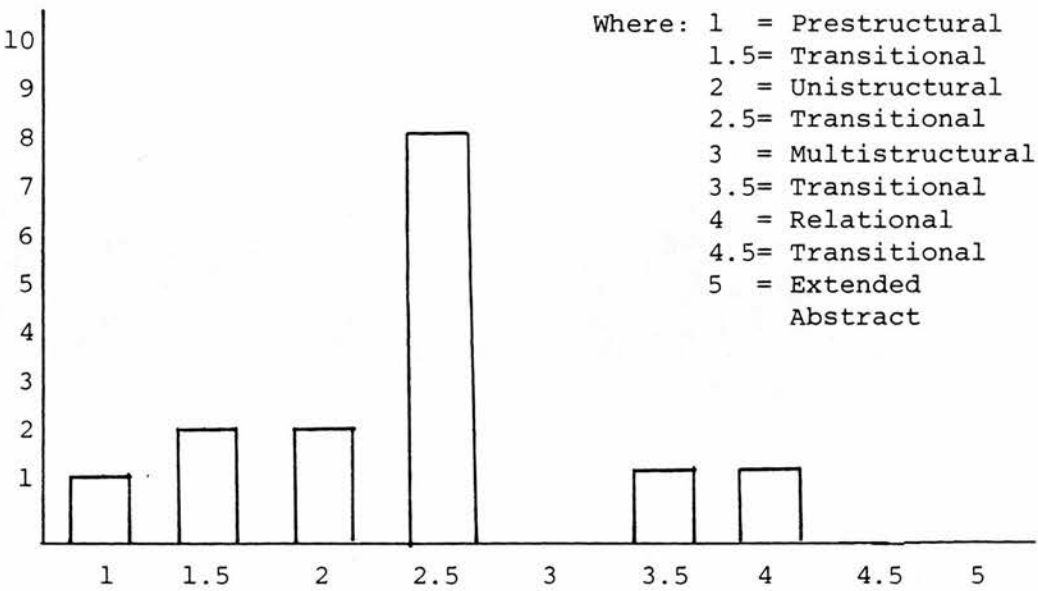


Figure 4.1 SOLO Scores: Beginners (n = 15), Test 1 (English).

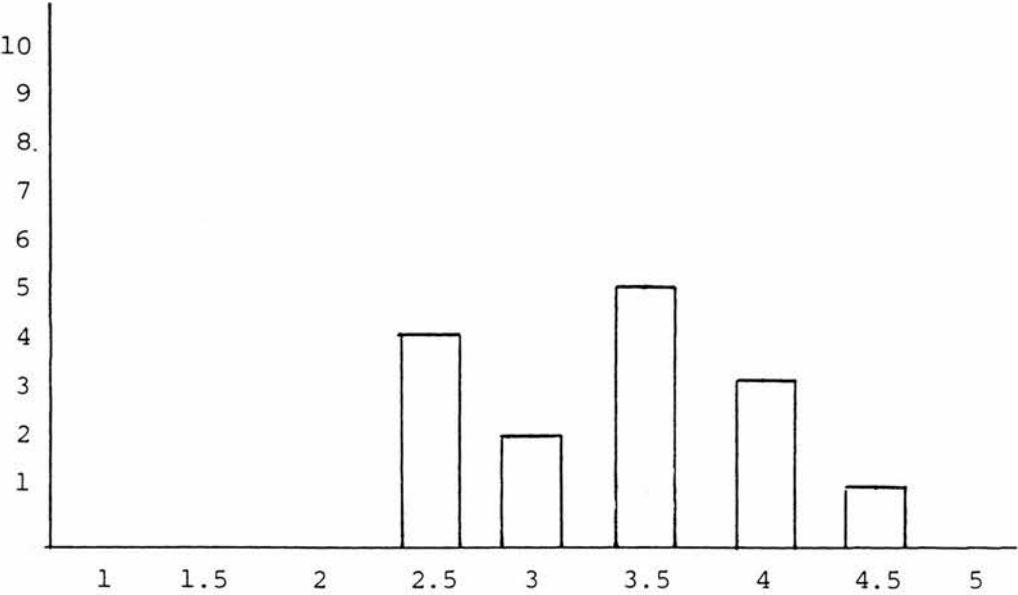


Figure 4.2 SOLO Scores: Intermediate (n = 15), Test 1 (English)

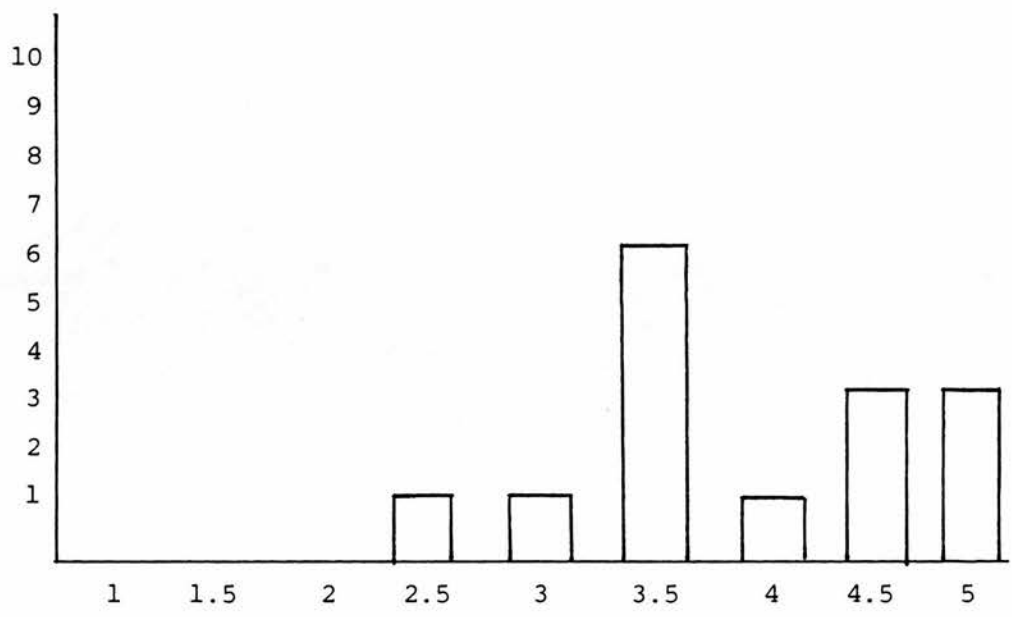


Figure 4.3 SOLO Scores: Advanced (n = 15) Test 1 (English)

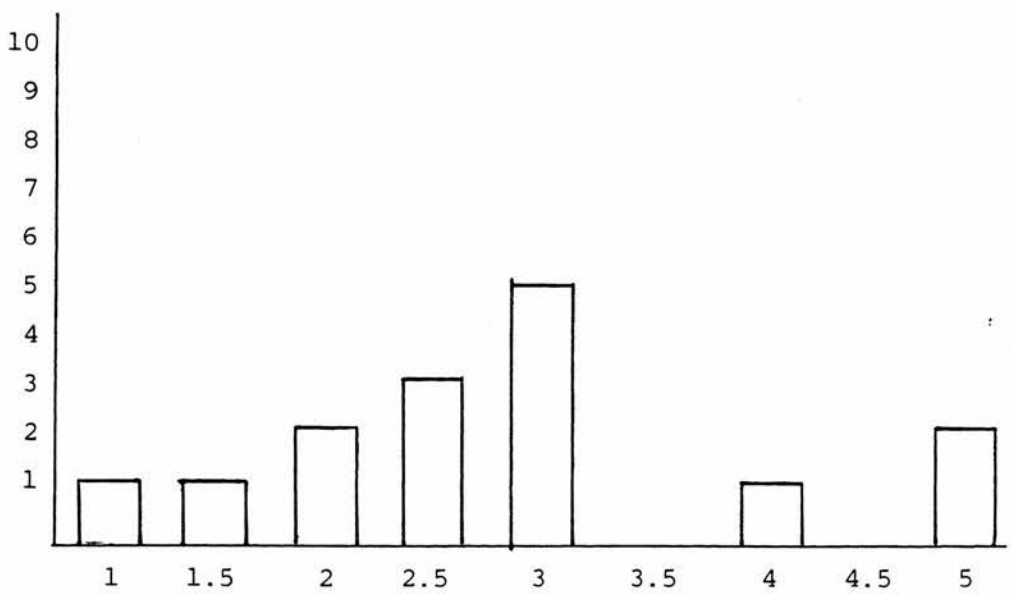


Figure 4.4 SOLO Scores: Beginners (n = 15) Test 2 (English)

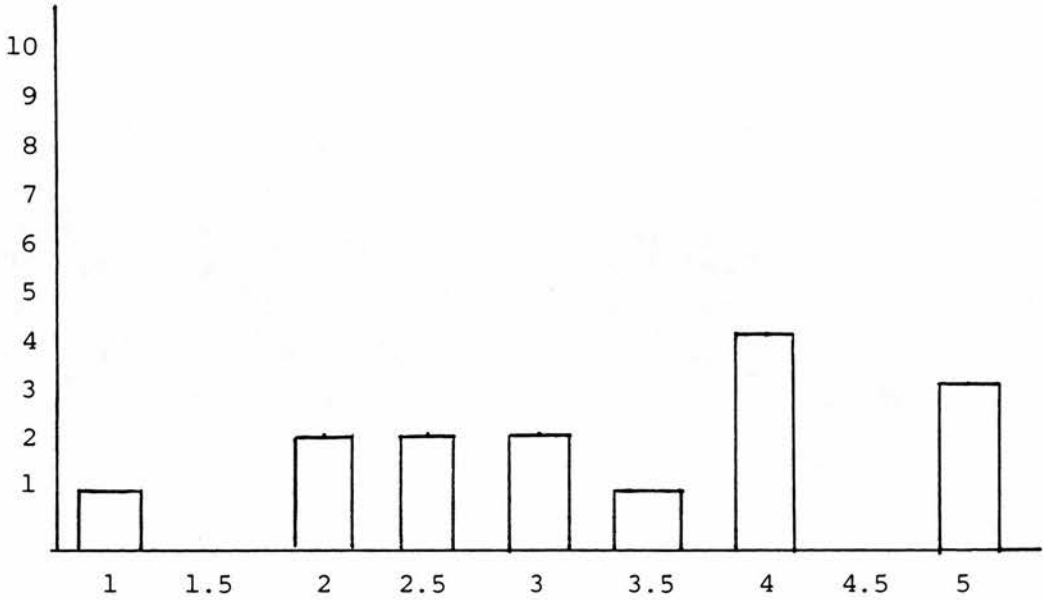


Figure 4.5 SOLO Scores: Intermediate (n = 15) Test 2 (English)

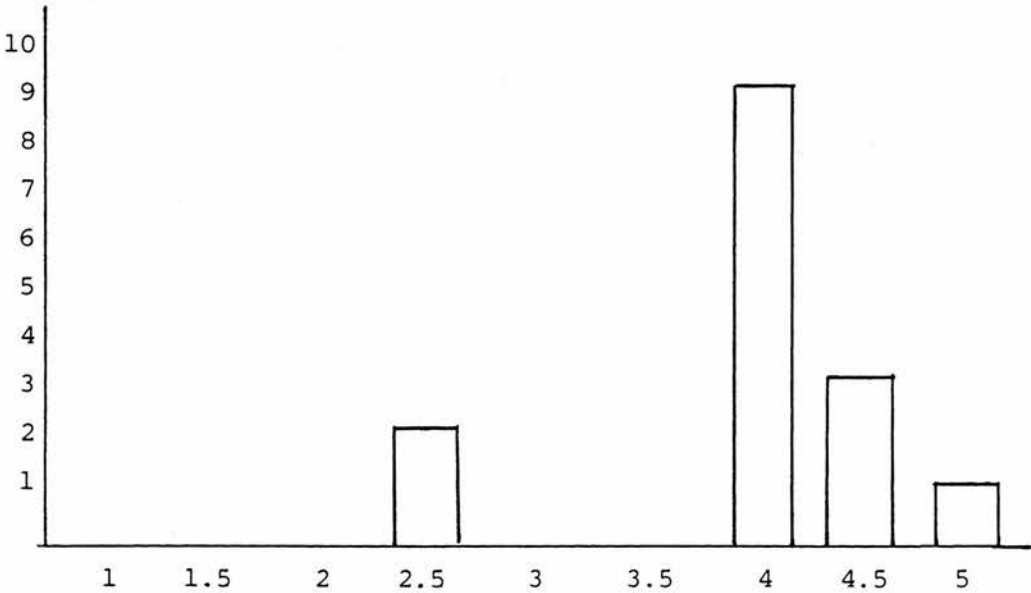


Figure 4.6 SOLO Scores: Advanced (n = 15) Test 2(English)

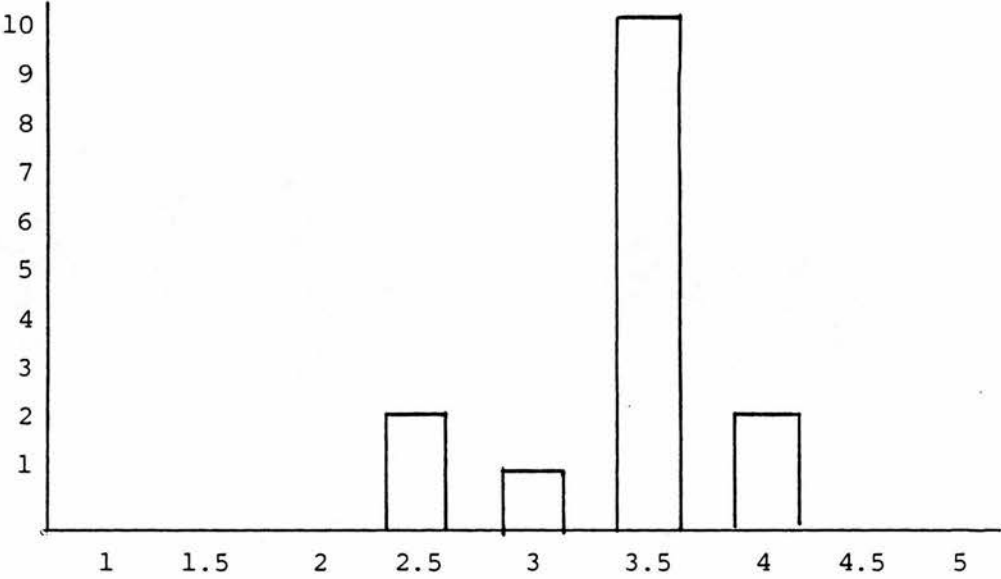


Figure 4.7 SOLO Scores: Beginners (n = 15) Test 1 (Spanish)

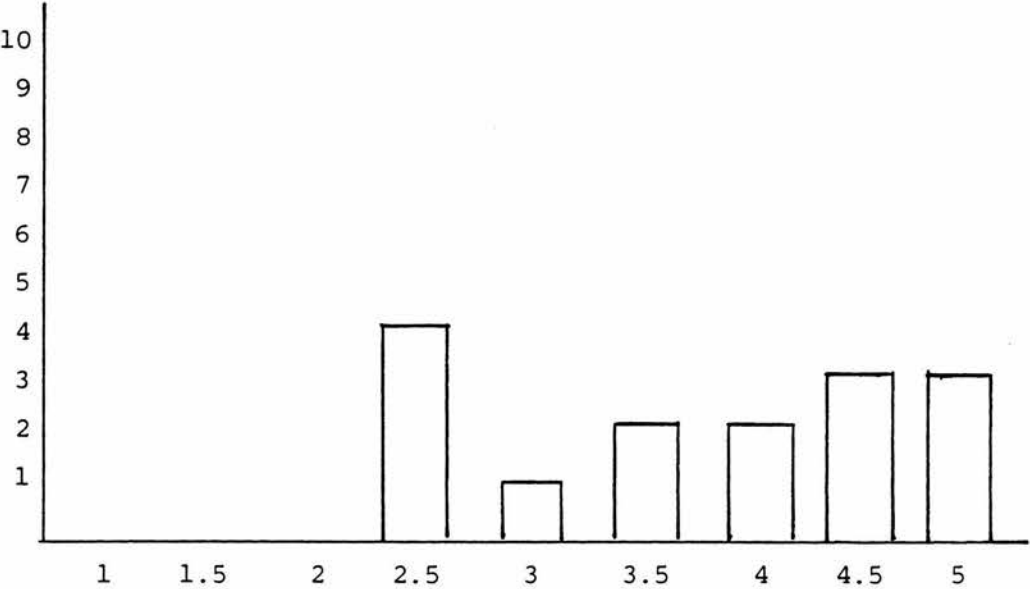


Figure 4.8 SOLO Scores: Intermediate (n = 15) Test 1 (Spanish)

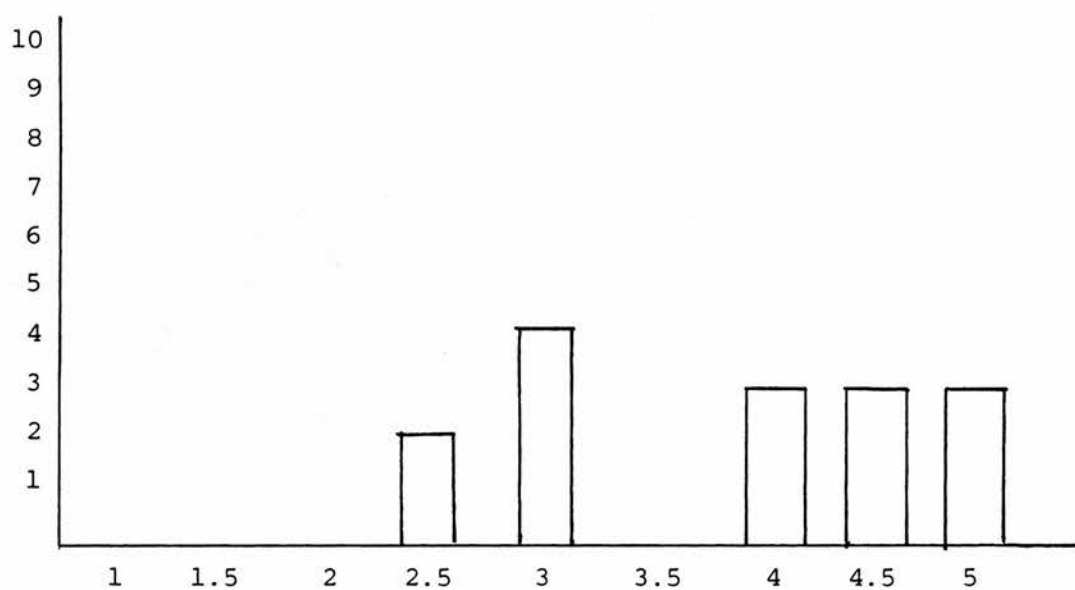


Figure 4.9 SOLO Scores: Advanced (n = 15) Test 1 (Spanish)

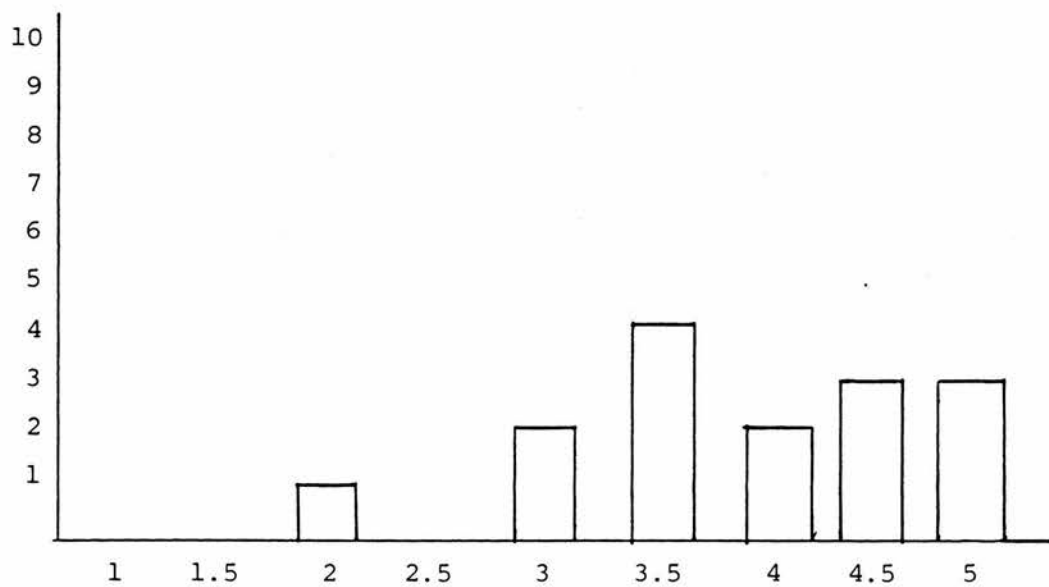


Figure 4.10 SOLO Scores: Beginners (n = 15) Test 2 (Spanish)

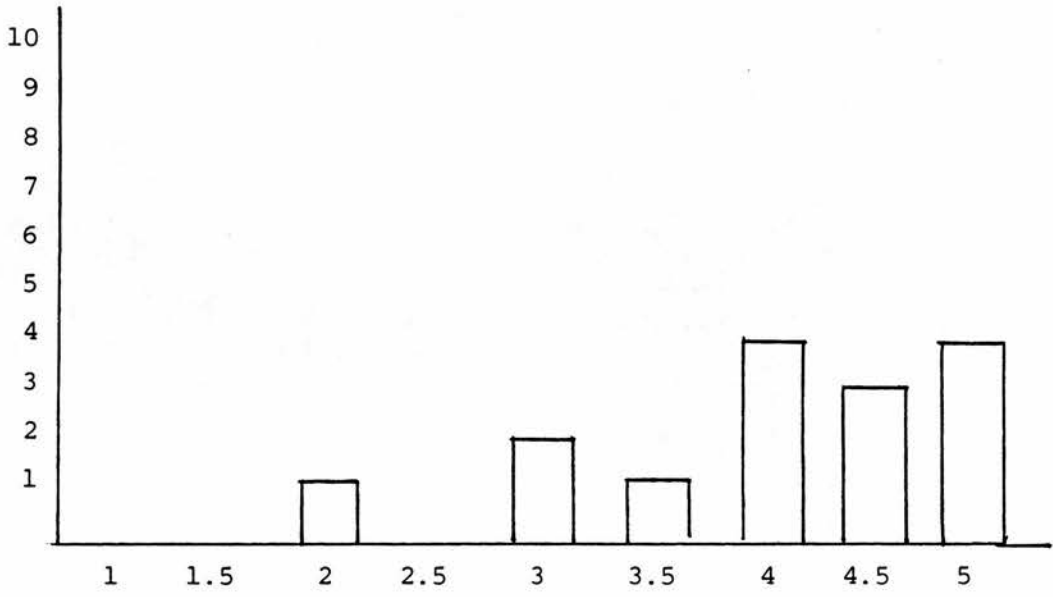


Figure 4.11 SOLO Scores: Intermediate (n = 15) Test 2 (Spanish)

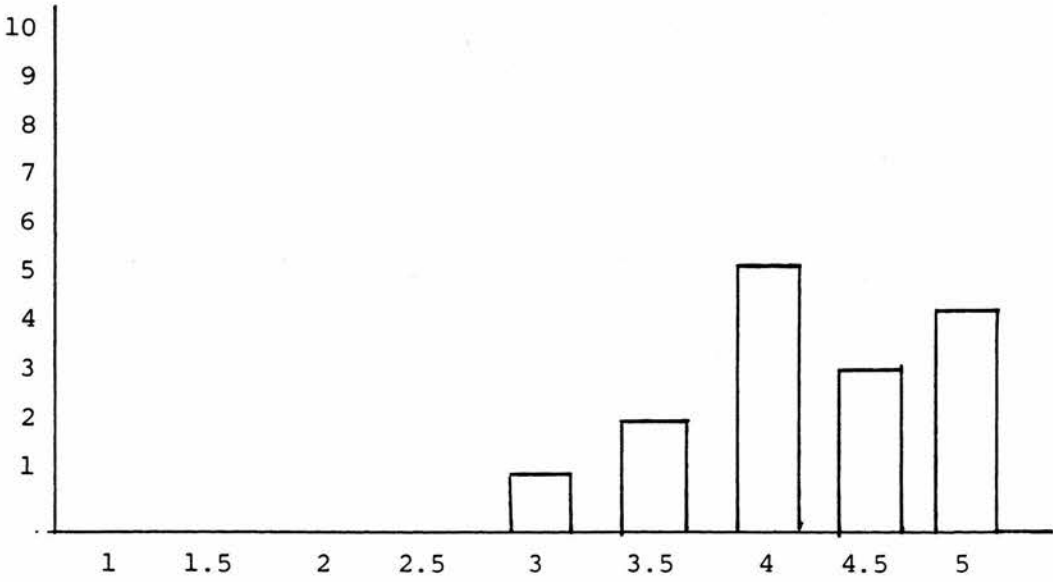


Figure 4.12 SOLO Scores: Advanced (n = 15) Test 2 (Spanish)

	ENGLISH						SPANISH						
	FIRST			SECOND			FIRST			SECOND			
SOLO Scores	B	I	A	B	I	A	B	I	A	B	I	A	TOTALS
1	1	0	0	1	1	0	0	0	0	0	0	0	3
1.5	2	0	0	1	0	0	0	0	0	0	0	0	3
2	2	0	0	2	2	0	0	0	0	1	1	0	8
2.5	8	4	1	3	2	2	2	4	2	0	0	0	28
3	0	2	1	5	2	0	1	1	4	2	2	1	21
3.5	1	5	6	0	1	0	10	2	0	4	1	2	32
4	1	3	1	1	4	9	2	2	3	2	4	5	37
4.5	0	1	3	0	0	3	0	3	3	3	3	3	22
5	0	0	3	2	3	1	0	3	3	3	4	4	26
TOTALS	15	15	15	15	15	15	15	15	15	15	15	15	180
Where: B= Beginners. I= Intermediate. A= Advanced.													

Table 4.6 SOLO Scores: Number of Subjects by "Language of Test," "Order of Administration" and "Level of Proficiency".

Already certain tendencies can be appreciated in the preceding histograms and table. However, to be able to make any assertions about them it was necessary to test the statistical significance of the differences observed among the different experimental conditions. The results of such test will be presented in Section 4.4, and discussed in Chapter 5.

Before presenting the results just mentioned, however, it will still be necessary to examine the degree of interjudge reliability observed. This will be done in the following section.

4.3.2 Interjudge Reliability

The first step to analyse the degree of agreement between the marks awarded by the two judges was to record the number of cases where there had been more than one level of difference in the initial scoring. This was compared to that of cases where there had been no difference, or the difference had been of up to one level only. Table 4.7 summarizes these data.

		T E S T		TOTALS
		ONE	TWO	
DIF•	None	22	26	48
FER•	Up to One Level	39	47	86
ENCE	More Than One Level	29	17	46
T O T A L S		90	90	180

Table 4.7 Differences Between Initial Scores Awarded by Both Judges in Both Tests.

The application of a χ^2 test to the above data showed that the difference in proportions was not significant at the 5% level, since the observed value of χ^2 was 3.3874, while the corresponding table value (for 2 d.f.) was 5.991 (Robson, 1973, Table G, p.151).

Although the proportion of cases where there was an initial disagreement between the two judges of more than one level did not prove to be significantly different from that of the other cases, such cases of disagreement were duly discussed. This discussion was very productive since it allowed a better understanding of the data and of the marking scheme.

On the other hand, the discussion also led to the modification of some scores. Once more, a χ^2 test helped to check if the difference in proportions between the modified and unmodified scores in Tests 1 and 2 was statistically significant. Table 4.8 shows the data used to compute the value of χ^2 .

	T E S T		TOTALS
	ONE	TWO	
MODIFIED	9	4	13
UNMODIFIED	20	13	33
TOTALS	29	17	46

Table 4.8 Number of Cases of Initial Disagreement of More than One Level: Modified and Unmodified.

Since the observed value of χ^2 for the above data was .0412 (with 1 d.f.), the corresponding table value was larger: 3.841 (Robson, op.cit., p.151). The proportion of cases where some modifications were made to the subjects' initial scores in Test 1 did not differ significantly from the corresponding one in Test 2.

Taking into consideration the results of these preliminary tests, it was decided to go on and examine the degree of agreement between the definitive sets of scores awarded to all subjects in both tests by Judges A and B. These scores are shown in Appendix L, Columns III and IV, respectively. The degree of agreement for each subject is shown in Column V of the same appendix, and Tables 4.9 and 4.10 specify the degree of agreement for Tests 1 and 2, respectively.

	Agree	Half Level Difference	One Level Differ- ence	One Level and a Half Dif- ference	Two Levels Difference	Total
N	25	26	18	19	2	90
Percentage	27.77	28.88	20	21.11	2.22	100

Table 4.9 Degree of Agreement Between Two Judges on Test 1.

	Agree	Half Level Difference	One Level Differ- ence	One Level and a Half Dif- ference	Two Levels Difference	Total
N	27	28	21	13	1	90
Percentage	30	31.11	23.33	14.44	1.11	100

Table 4.10 Degree of Agreement Between Two Judges on Test 2.

It will be noticed that the pattern is very similar in both tests. In both, the proportion of cases where there was a difference of half level only is the predominant one (28.88%, and 31.11%, respectively). This is immediately followed by the proportion of cases where there was complete agreement (27.77%, and 28.88%). The groups that follow are those where there was a difference of one level and a half (21.11%), and one level (20%), in Test 1. In Test 2, the order is reversed: the group where there was a difference of one level (23.33%) is followed by the one where the difference was of one level and a half (14.44%). In both tests, the smallest group is that where there was a two-level difference (2.22% and 1.11%). It can be noticed, finally, that in none of the tests was a difference of more than two levels recorded. Table 4.11 specifies the degree of agreement between the two judges taking into consideration both tests at the same time.

	Agree	Half Level Difference	One Level Differ- ence	One Level and a Half Dif- ference	Two Levels Difference	Total
N	52	54	39	32	3	180
Percentage	28.88	30	21.66	17.77	1.66	100

Table 4.11 Degree of Agreement Between Two Judges on Tests 1 and 2.

As a next step, the Pearson product - moment correlation coefficient was calculated for the definitive sets of scores awarded by Judges A and B. The correlation coefficient was worked out for each

of the tests first, and then for both of them together (Table 4.12). The statistical significance of the observed values was determined by checking the corresponding table values for two-tailed tests (Hatch and Farhady, 1982, p.277).

Test	d.f.	r	p
Test 1	88	.6804	<.001
Test 2	88	.7260	<.001
Tests 1 and 2	178	.6879	<.001

Table 4.12 Correlation Between Two Judges on Test 1, Test 2, and Tests 1 and 2.

As can be seen in Table 4.12, the probability of having found these coefficients by chance is very low. This fact, supported by the analyses previously made of the differences between the two judges in both tests (Tables 4.9, 4.10 and 4.11), can be taken as evidence of a high degree of interjudge reliability of the marking scheme used.

Having made sure that the testing instrument used satisfied the requirement of interjudge reliability, it was considered appropriate to submit the data to the intended ANOVA test. Its results will be presented in the next section.

4.4 ANOVA

It has been pointed out before (Sections 3.2.2 and 4.1) that owing to the fact that the experimental design of the study is a three-way factorial design, a three-way ANOVA test was considered appropriate to determine if the differences observed amongst the contrasting groups were significant or not.

Furthermore, Norusis states regarding the proper application of the ANOVA test that "the dependent variable must be interval level, and one or more categorical variables define the groups." (Norusis, 1985, p.426). In the present study this is indeed the case, since the independent variables are categorical: "Language of Test" (English/Spanish), "Level of Proficiency" (Beginners/Intermediate/Advanced), and "Order of Administration" (First/Second); and the dependent variable ("Reading Outcome"), interval.

Once the decision to use this test was taken, it was necessary to choose a specific model suitable for this study, for as Iversen and Norpoth note in their discussion of models of analysis of variance:

As an investigator considers an increasing number of explanatory variables, he is faced with the problem of how to observe behavior in increasingly complex circumstances. In order to make the gathering less cumbersome and less costly, and perhaps at all feasible, the investigator can avail himself/herself of several designs...

(Iversen and Norpoth, 1976, p.83).

In this respect, Bray and Maxwell - in their discussion of the use of ANOVA, MANOVA and related techniques in the behavioural sciences - point out that "there is no one 'right' method but instead...there are advantages and disadvantages to each, which depend on what the researcher hopes to glean from the data."

(Bray and Maxwell, 1985, p.8). The approach taken in their discussion is to "focus on the appropriate use and interpretation of the various methods." (Ibid., p.9).

A similar position is taken by Gaito regarding analysis of variance techniques in repeated measurements designs. He points out that "the overall problem of repeated measurements designs is a complex one, and a satisfactory treatment has not been effected." (Gaito, 1970, p.289). In his discussion of the advantages and disadvantages of six possibilities he presents for dealing with this problem, he states that "it would appear that if one does have a repeated measurements design, the safest procedure would be to randomize the order of treatments..." (Ibid., p.295). This, it will be remembered, was done in the present study (cf. Section 3.4).

Let us now consider the procedure chosen for this study. It will be convenient to start by pointing out that Kerlinger's equation shown in Figure 4.13 provided the basic framework for the procedure used.

$V_t = V_b + V_i + V_e$	
Where:	V_t = total variance V_b = between-groups variance V_i = variance due to individual differences V_e = "true" error variance

Figure 4.13 Analysis of Variance:
Kerlinger's Equation Used
in this Study (Kerlinger,
1967, p.231).

In his discussion of this equation, Kerlinger points out that the within-groups variance contains more than variance error, i.e., that it also contains variance due to individual differences, and that this equation captures this fact. He also states that "if we can find a way to control or measure V_i , to separate from V_w , then it follows that a more accurate measure of the 'true' error variance is possible." (Ibid.)

Let us go back to the procedure followed in this study. In the first place, a straightforward analysis of variance was carried out, following a procedure recommended by Moroney for factorial designs (Moroney, 1951, pp. 394-420). Then, the variance due to individual differences within groups was estimated following a procedure based on examples given by Hays (1973, pp.551-74). This source of variation was separated from the within-groups variance that had been found following Moroney's procedure. The resulting Table of Analysis of Variance will be presented in the following sections, along with other relevant information.

Finally, following Moroney's advice (op.cit., pp.415-6), two breakdown analyses of variance were carried out to further analyse the interaction "Language of Test" x "Level of Proficiency," which was the only interaction that reached significance in the previous step. The results of these breakdown analyses by "Language of Test" and by "Level of Proficiency" will be presented in Section 4.4.2.

Before going on to consider the results of the complete analysis of variance, I would only like to point out that the averages of the highest scores awarded to each subject by Judges A and B provided the data for these analyses. As stated before (cf. Section 4.3.1), averages were considered appropriate for these

analyses of the data, since they capture both judges' evaluation of each subject. The averages in question can be found in Appendix L, Column VII.

4.4.1 Complete ANOVA

The results of the application of the ANOVA test following the procedure described in the preceding section will be examined by making reference to the corresponding Table of Analysis of Variance (Table 4.13).

However, before doing so, some information regarding the tables used - following Moroney - in the successive steps of the calculations resulting in the Table of Analysis of Variance may be necessary. These tables are included in Appendix M (Tables M.1 through M.5).

The first one of such tables (Table M.1) contains the original data from which the following ones were compiled. The data included in this table has been arranged as follows: the scores in the columns labelled "First" have been rank ordered, both for the English and the Spanish tests. Then, the scores obtained by the subjects in their second test has been recorded in the same row, in the corresponding "Second" columns. It must be noticed, then, that the scores in the column "English-First," and those in the column "Spanish-Second," correspond to the same subjects. That is why these columns are marked with an asterisk (*). Likewise, the scores of the subjects who took the Spanish test first, and the English test second, are included in the corresponding columns: "Spanish-First," and "English-Second." To identify them, these

two columns have been marked with two asterisks (**). So, for example, the first subject in the group of "Beginners" - under "English-First" - scored 3 points in the English test, which he took first, and 5 points in the Spanish test, which he took second (cf. same row, column "Spanish-Second"). The next subject in that group scored 2.75 in the English test, and 4 in the Spanish one. The third one scored 2.5 in both tests, and so on.

On the other hand, the subject in the group of "Intermediate" students who scored 5 points in the Spanish test, having taken it first (cf. the "Spanish-First" column), scored 4 points in the English test (cf. the "English-Second" column). The second subject in this group scored 4.75 points in the Spanish test, and 2 points in the English one, and so on.

This specific way of arranging the data in Table M.1 was chosen to facilitate the compilation of Table M.5, which was used to estimate the variance due to individual differences to be separated from the within-groups variance.

In this respect, it may be convenient to note that the corresponding between-subjects table value (i.e., that one due to variation between individuals within groups) was estimated - as already noted in the preceding section - following examples given by Hays (1973, pp.551-74). However, a formula based on those used by Lewis (1967, p.148) could also be applied (Figure 4.14), yielding the same results.

$$S_s = S \frac{(\text{Level Sum})^2}{(\text{Level Size})} - \frac{(S_x)^2}{N}$$

where:

S_s = Between-subjects sum of squares

Level = Cell

$\frac{(S_x)^2}{N}$ = Correction factor

d.f. = Level Size - 1

Figure 4.14 Equation to Obtain the
Between-Subjects Sum of Squares.

To apply this formula to the data in this study, one only has to divide the sum of squares of the entries in Table M.5 by the number of items (repeated measures) going to make each entry (i.e., 15), and then subtract the correction factor. Between the 15 items in each cell (level) there are 14 d.f.

The other tables in Appendix M (Tables M.2 through M.4.3) are also based on the information included in Table M.1, and were compiled following the steps in the procedure suggested by Moroney (op.cit., pp. 400-11).

Let us now turn our attention to the resulting Table of Analysis of Variance (Table 4.13):

SOURCE	SUM OF SQUARES	d.f.	VARIANCE ESTIMATE	<u>F</u>
Language (of Test)	16.0503	1	16.0503	38.9098
Level (of Proficiency)	24.7340	2	12.3670	29.9806
Order (of Administration)	5.2534	1	5.2534	12.7355***
Language x Level	6.1094	2	3.0547	7.4053***
Order x Level	1.3393	2	.6696	1.6232 (n.s.)
Language x Order	0.1532	1	.1532	
Language x Order x Level	0.8302	2	.4154	1.0070 (n.s.)
Residual	63.5313	154	.4125	
Total	118.0011	165		
Between Subjects	51.4368	14		
Grand Total	169.4379	179		

Table 4.13 Table of Analysis of Variance (Complete ANOVA).

The first thing to explain about this table is the way in which the "Grand Total" of the "Sum of Squares" - arrived at following Moroney's procedure (op.cit., pp.394-412) - has been partitioned. As pointed out in the preceding section, this has been done in agreement with the partitioning of the variance in Kerlinger's equation: $V_t = V_b + V_i + V_e$ (cf. Figure 4.13). That is, the "Between Subjects" sum of squares (i.e., the one due to variation between individuals within groups) has been separated from the "Grand Total," leaving a "Total" of 118.0011. From this, the "explained" sums of squares (i.e., those due to the main effects and to the interactions; in other words, the between-treatments sums of squares) have been subtracted, leaving a "Residual" ("error") sum of squares of 63.5313.

The partitioning of the "Grand Total" number of degrees of freedom has been performed in the same way: from the 179 d.f. in the "Grand Total," the 14 "Between Subjects" degrees of freedom have been subtracted. From the remaining "Total" (165 d.f.), those corresponding to the main factors and to the interaction (11 d.f.) have been removed, leaving 154 degrees of freedom for the "Residual."

The figures in the "Variance Estimate" column resulted from dividing the "Sum of Squares" associated with each source of variation by its corresponding number of degrees of freedom. Thus, for example, the sum of squares associated with the main factor "Language" (of Test): 16.0503 was divided by the appropriate number of degrees of freedom (1), resulting in a "Variance Estimate" of 16.0503.

The next step was to test whether the variance estimate associated with each of the main effects and interactions was significantly greater than that of the residual (F test). The F ratio for each main effect or interaction was therefore calculated by dividing its variance estimate by the variance estimate associated with the residual (e.g., 16.0503 divided by .4125 = 38.9098).

Following Moroney's advice op.cit., pp. 413-8), the F test was applied to the second order interaction first ("Language x Order x Level"). The variance estimate associated with this interaction is just slightly more than the residual variance estimate. It is therefore not significantly greater, hence the (n.s.) in front of the corresponding value of F (1.0070) in the table.

Moving upward to the first order interactions, the F test was not applied to the "Language x Order" interaction, since the variance estimate associated with it (.1532) is smaller than that of the residual (.4125). There was no point, therefore, in trying to test whether it was "significantly greater." This explains the blank that has been left in the space corresponding to this interaction in the F column.

Moving up to the "Order x Level" interaction, it can be noticed that the variance estimate associated with it is greater than the residual variance estimate. However, the difference is not large enough to be significant.

On the other hand, in the case of the "Language x Level" interaction, it was found that the variance ratio (F = 7.4053), for the corresponding 2 and 154 degrees of freedom, was indeed significant

beyond the 0.1% level, since the table value of \underline{F} for 2 and ∞ degrees of freedom at the .001 level is 6.91 (Table 3, Lindquist, 1953, p.44); hence the three asterisks in front of the observed value of \underline{F} for this interaction in the table.

This means that the "Language" and "Level" effects are not independent on their effect on the SOLO scores of the subjects. Some combinations are better or worse than we should expect if one expected the null hypothesis (of no difference) to be true.

Having found that this interaction reached significance, it was necessary to do - as Moroney points out (op.cit., p.415) - repeated analyses of variance taking the interaction into account. The repeated (or breakdown) analyses of variance that were run on the data to understand the way in which the explanatory variables "Language" (of Test) and "Level" (of Proficiency) interacted in this study will be reported in the following section.

For the moment, however, let us go back to the table and move up to the main effects. The first one to be considered, then, is "Order" (of Administration). This main effect, it must be pointed out, is the one which was not involved in the only first order interaction which was significant. It was, therefore, appropriate to test if its variance estimate was significantly greater than that of the residual (cf. Moroney, op.cit., p.413).

As a result of the \underline{F} test for this effect, it was found that it was significant beyond the 0.1% level. Its observed value was 12.7355 (with 1 and 154 degrees of freedom), and the table value of \underline{F} for $1/\infty$ d.f. was 10.83 at the 0.1% level (Lindquist, 1953, p.44). This explains, as in the case of the interaction

"Language x Level," the three asterisks that follow the value of F in the table.

Although the implications of having found this main effect significant will be taken up in a later chapter, the following graphs (Figures 4.15 and 4.16) can be useful at this stage to illustrate the way in which this explanatory variable affected the subjects' reading outcomes (the dependent variable). I will use two graphs, following Spector (1981, pp.65-6).

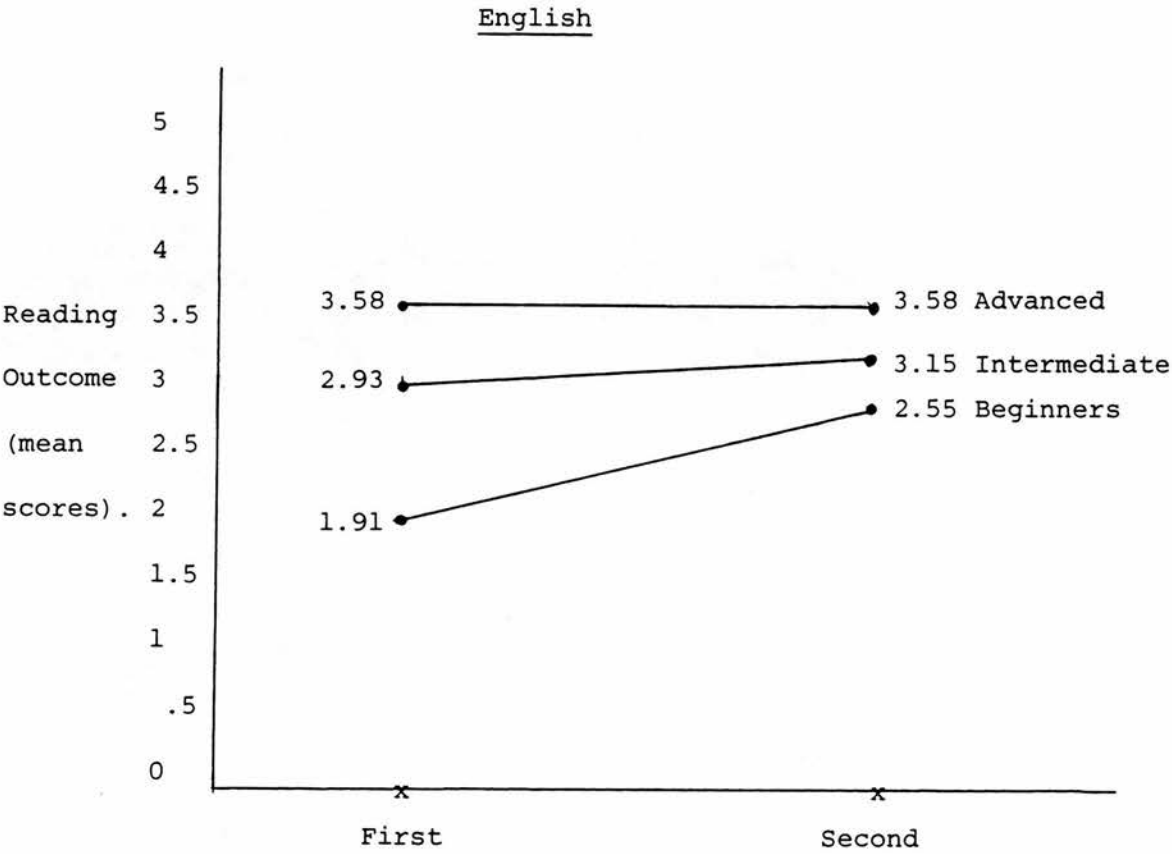


Figure 4.15 Graphic Representation of the "Order"
(of Adminstration) Main Effect:
English Tests.

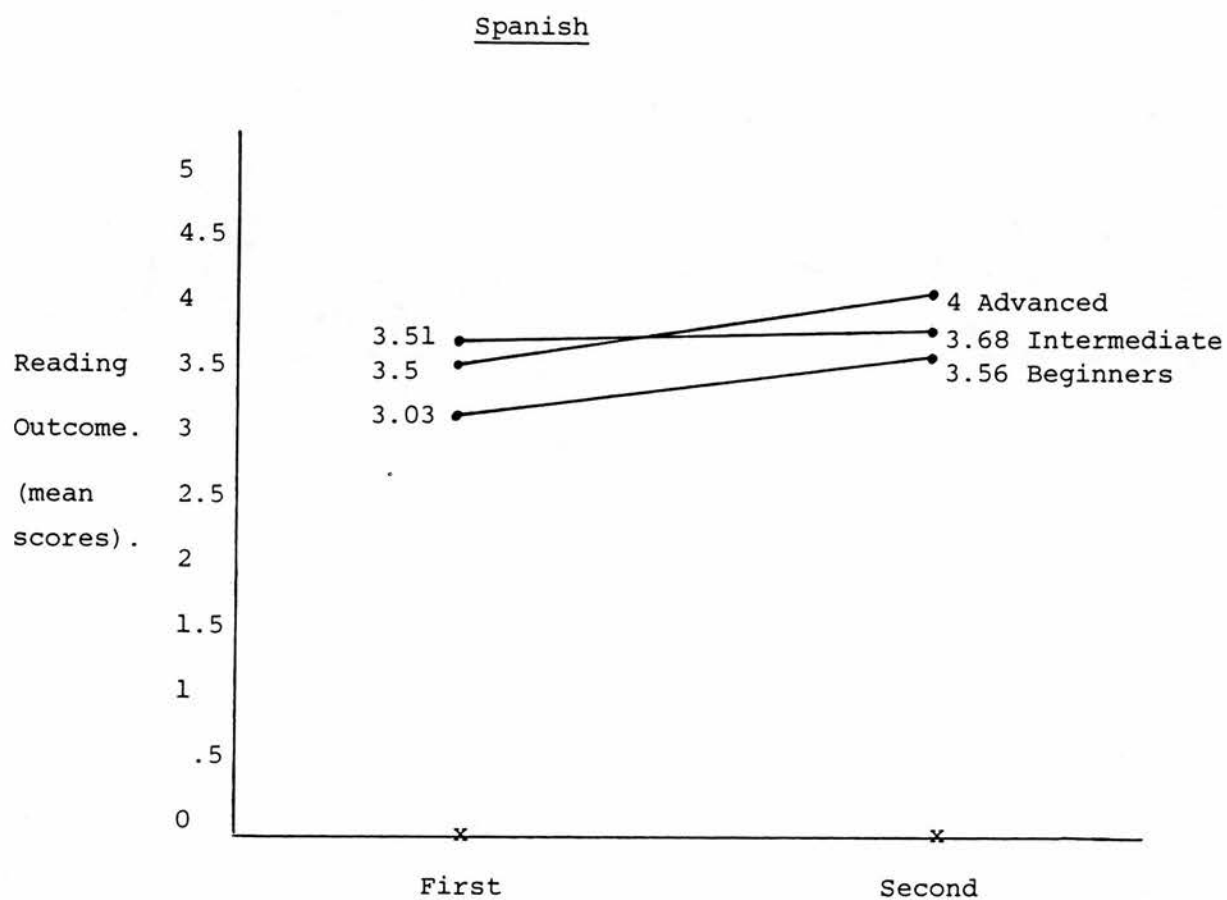


Figure 4.16 Graphic Representation of the "Order"
(of Administration) Main Effect:
Spanish Tests.

The first graph (Figure 4.15) corresponds to the results in the English tests, and the second one (Figure 4.16) to those of the Spanish tests. The y-axis represents the dependent variable ("Reading Outcome"). The x-axis represents the independent variable under consideration: "Order" (of Administration), which takes on its two values: "First" and "Second." The "Level" (of Proficiency) variable is represented by three lines ("Beginners," "Intermediate" and "Advanced"). The values of the dependent variable used in these graphs are the means corresponding to each cell in Table M.2 (Appendix M), but it should be noticed that the figures shown in that table could have been used instead, with the same graphic results.

It will be noticed that except for the case of the two groups of "Advanced" subjects, whose performance in the "English" test was the same regardless of whether it was administered to them as the "First" or the "Second" test (cf. Figure 4.15), all the other groups who took the "English" or "Spanish" tests in "Second" order outperformed those who took them first. For example, the "Advanced" subjects who took the "Spanish" test in "Second" order outperformed those "Advanced" subjects who took the "Spanish" test in the "First" order (cf. Figure 4.16), and so on.

Let us go back to the Table of Analysis of Variance (Table 4.13) once more, and consider the two remaining main effects: "Language" (of Test), and "Level" (of Proficiency).

It must be pointed out that no F test was applied to them, since they were involved in the only first order interaction that was significant ("Language x Level"). As Moroney indicates, "if a

first order interaction were to prove significant, then we should not be justified in testing any of the main factor effects related to that interaction against the residual." (Op.cit., p.413). That is why no asterisks appear in front of their corresponding F ratios in spite of the fact that they are even larger than that corresponding to the "Order" effect.

Finally, it should be remembered that a fuller account of the way in which "Language" (of Test) and "Level" (of Proficiency) interact upon the dependent variable will be given in the following section where the corresponding breakdown analyses are presented.

4.4.2 Breakdown Analyses of Variance

Two breakdown analyses of variance were carried out to examine the interaction "Language x Level," which was the only interaction that proved significant in the complete analysis of variance reported in the preceding section (cf. Table 4.13).

The results of the breakdown analyses will be presented separately. However, it may be convenient to illustrate graphically the interaction in question first, in order to get a clearer perspective for the subsequent analyses. The graphs in Figures 4.17 and 4.18 will serve for this purpose.

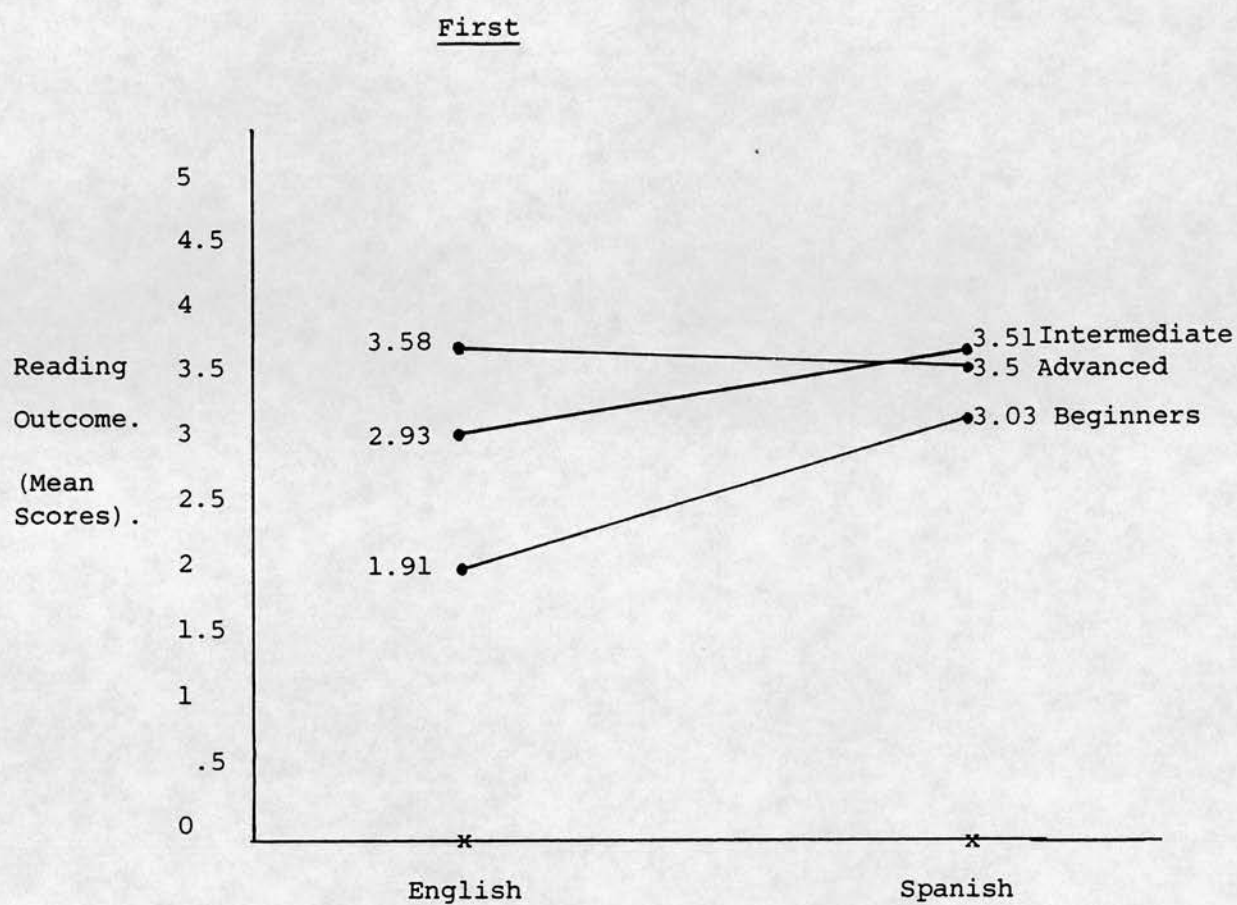


Figure 4.17 Graphic Representation of the Interaction "Language x Level." (Results of the First Administration of the Tests).

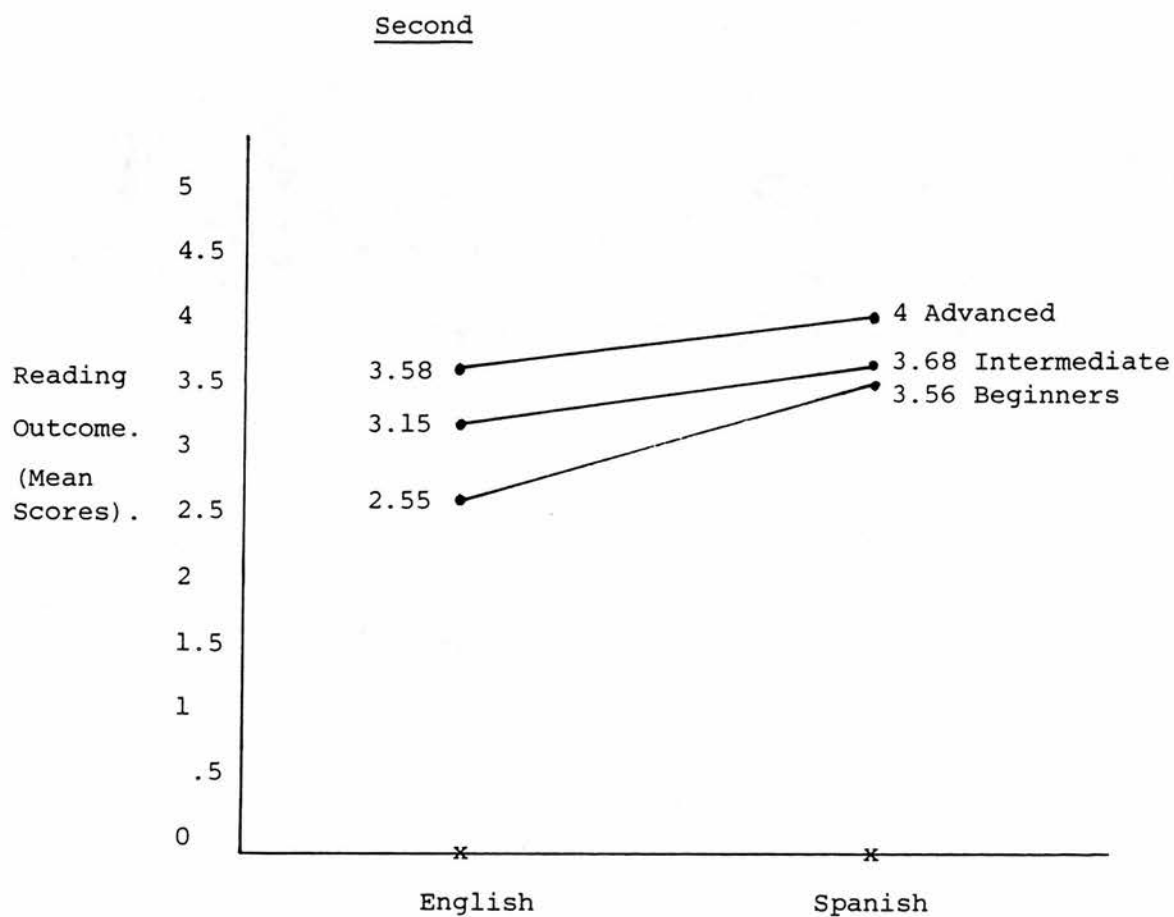


Figure 4.18 Graphic Representation of the Interaction
"Language x Level." (Results of the Second
Administration of the Tests).

The first graph (Figure 4.17) corresponds to the results recorded for the first administration of the tests, and the second graph (Figure 4.18) to those of the second administration. As is the convention for these graphs, the y-axis represents the dependent variable ("Reading Outcome"). The x-axis represents in both graphs the "Language" (of Test) variable, which takes on its two values: "English" and "Spanish." The "Level" (of Proficiency) variable is represented by the three lines ("Beginners," "Intermediate" and "Advanced"). The values of the dependent variable in these graphs are the means for each of the experimental conditions in question.

It can be seen in these graphs that there is an interaction of the explanatory variables "Language" and "Level" because the lines representing the "Level" (of Proficiency) variable are nonparallel.

As Iversen and Norpoth (1976, pp.72-3), and Moroney (op.cit., p.415), point out, if there had been no interaction the lines would have been parallel. For the sake of comparison, this hypothetical situation would have been represented by graphs like those in Figures 4.19 and 4.20. No differential effect would have been noticed amongst the contrasting groups.

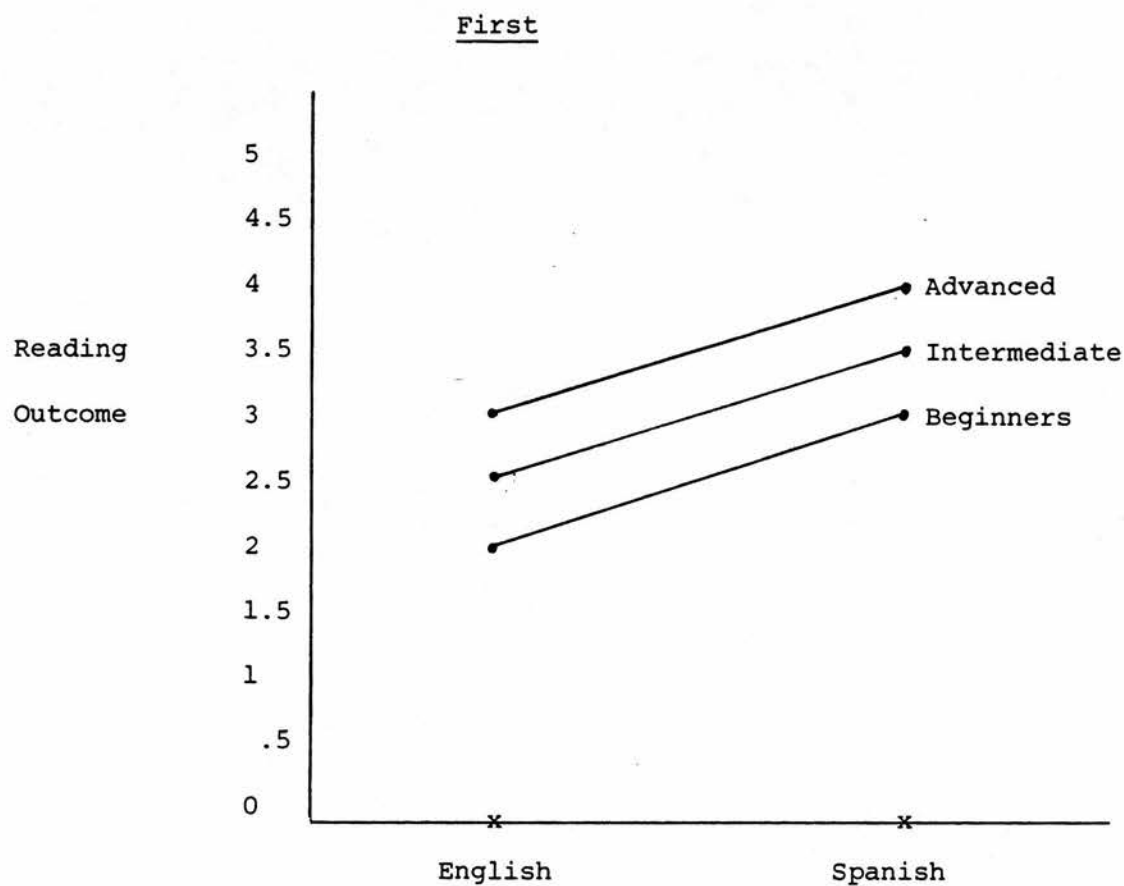


Figure 4.19 Graphic Representation of Lack of Interaction Between the Variables "Language" (of Test) and "Level" (of Proficiency): First Administration (Hypothetical).

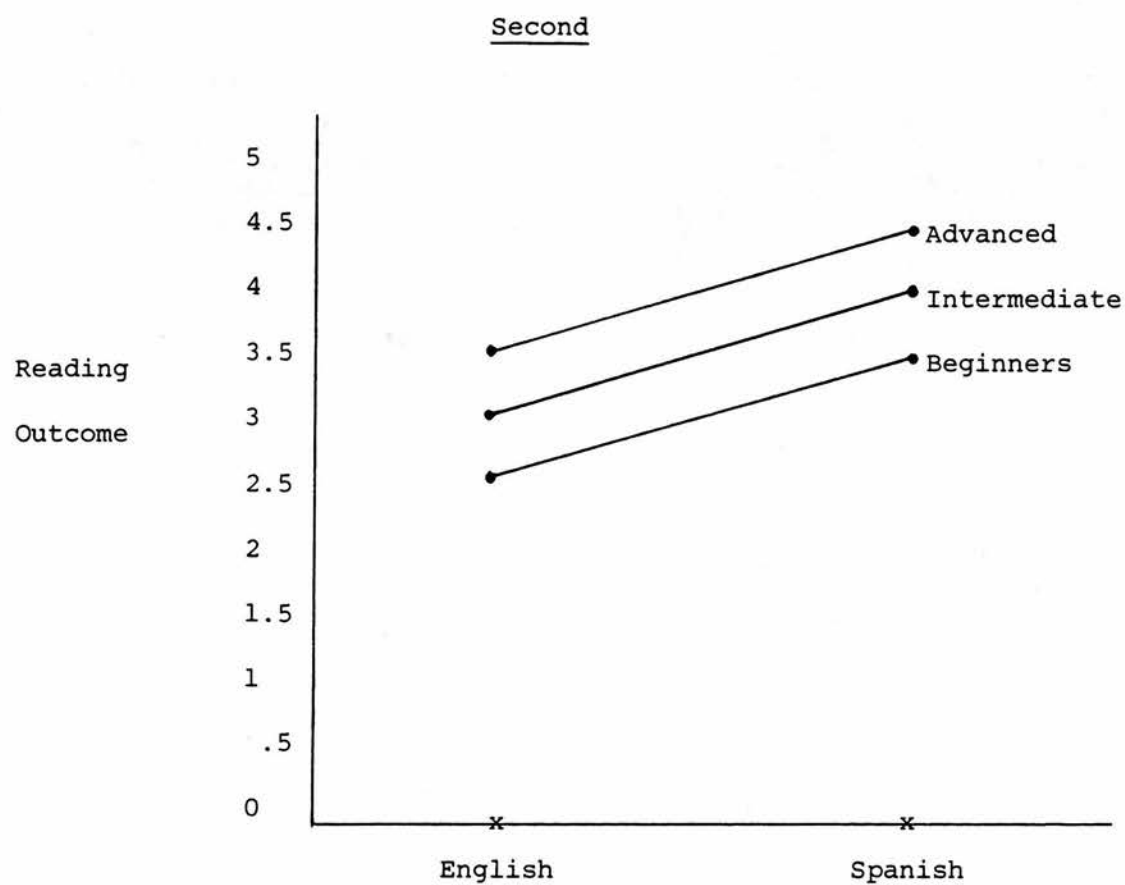


Figure 4.20 Graphic Representation of Lack of Interaction Between the Variables "Language" (of Test) and "Level" (of Proficiency): Second Administration. (Hypothetical).

Going back to the actual results, it can be noticed that in one occasion the lines representing the "Advanced" and the "Intermediate" levels even crossed (in the results of the first administration of the tests), as can be seen in Figure 4.17. This was due to the fact that the group of "Advanced" subjects who took the "English" test first outperformed those "Advanced" subjects who took the "Spanish" test first, whereas the reverse pattern was observed in the case of the "Intermediate" groups. Notice, too, that while this was also the case for the "Beginners" groups, the difference between the "Beginners" who took the "English" test first and those who took the "Spanish" test first is larger than that observed in the case of the "Intermediate" groups. In short, differential effects of the two variables were observed in the experimental groups. A similar observation can be made regarding the results depicted in Figure 4.18.

Let us now examine the results of the breakdown analyses of variance performed on the data. The tables used in these analyses are those included in Appendix N. The procedure followed to arrive at the respective Tables of Analysis of Variance was basically the same as that followed in the complete ANOVA test previously reported, the main difference being that fewer variables are being considered at a time now.

4.4.2.1 Breakdown Analysis of Variance by "Language" (of Test)

Actually, five separate analyses were carried out on the original data taking into account the interaction "Language x Level." In this section consideration will be given to two, i.e., those by "Language" ("English" and "Spanish"), and in the next section, to the three by "Level" ("Advanced," "Intermediate" and "Beginners").

Moroney (op.cit., p.416) points out that "conclusions arrived at in the sectional analyses of a breakdown analysis are valid for the level of the variable for which they are carried out." In this case, they will be valid for the "English" or "Spanish" levels of the variable "Language."

So, the first sectional analysis of the breakdown analysis by "Language" will be that one made for all the entries in Table M.1 (Appendix M) which were recorded at the level "English." The second one will be made for all the entries recorded in the same table at the level "Spanish."

Each of the analyses is a two-factor analysis: "Language" (of Test) x "Level" (of Proficiency). Since the "Language" variable has two levels ("English First," and "English Second"; or "Spanish First" and "Spanish Second"), and the "Level" (of Proficiency) variable has three ("Advanced," "Intermediate" and "Beginners"), the tests used were 2 x 3 ANOVA's in both cases.

There are, then, two Tables of Analysis of Variance to be considered in this section ("English" and "Spanish"). They are based on the information included in Tables N.1 (English) through N.3 (English), and N.1 (Spanish) through N.3 (Spanish), in Appendix N, respectively.

Let us take a look at the "English" one first (Table 4.14).

Source	Sum of Squares	d.f.	Variance Estimate	\underline{F}
"English"	1.806	1	1.806	2.3964 (n.s.)
"Level"	27.6937	2	13.8468	18.374***
"English x Level"	1.5544	2	.777	1.031 (n.s.)
Residual	63.3079	84	.7536	
Total	94.3625	89		

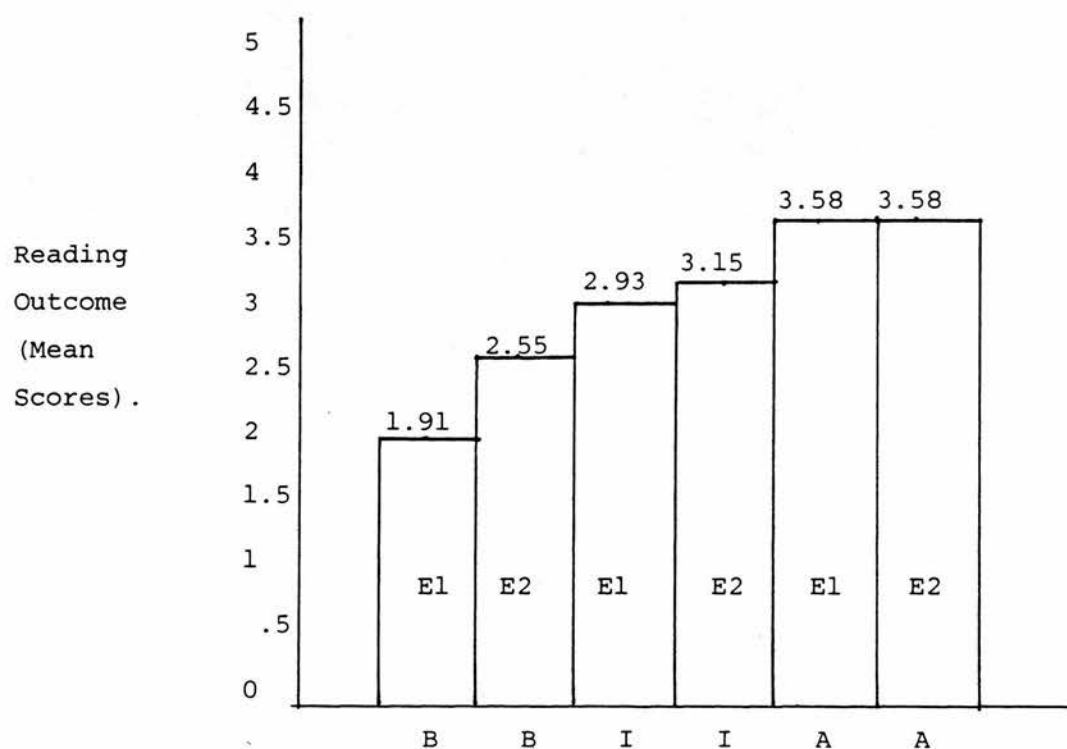
Table 4.14 Table of Analysis of Variance. Breakdown ANOVA by "Language": "English."

It will be noticed that the variance estimate associated with the "English x Level" interaction is larger than that associated with the "Residual." However, the \underline{F} ratio (1.031) was not significant. This means that there was not a significant interaction between the levels of the two variables in question. That is, the "English First," and "English Second" levels did not interact with the "Advanced," "Intermediate" and "Beginners" ones. No differential effect of the order in which the English test was administered was found regarding the three groups formed according to the subjects' degree of command of that language.

Moving upwards in the table, it can be seen that in the case of the main factor "Level," the variance ratio ($\underline{F} = 18.374$) for 2 and 84 degrees of freedom was significant beyond the 0.1% level, the table value of \underline{F} being 7.76 for 2 and 60 degrees of freedom (Table 3, Lindquist, 1953, p.44). In other words, "Level" of proficiency matters: the more advanced the higher the results.

Finally, the \underline{F} value for the other main effect ("English"), was not significant. That is, having taken the English test in first ("English First") or second place ("English Second") did not result in a significant difference.

Figure 4.21 illustrates these results. Since this time only two variables are involved, it is possible to include all of the relative information in one graph only. These results - as indeed those of all the other analyses performed - will be taken up in a later chapter, in the overall discussion of the results.



Where:

E1 = "English First"
 E2 = "English Second"
 B = "Beginners"
 I = "Intermediate"
 A = "Advanced"

Figure 4.21 Reading Outcome as Determined by the "Beginners," "Intermediate" and "Advanced" Levels of the "Level" (of Proficiency) Variable, and by the "English First" and "English Second" Levels of the "English" Variable.

Let us now consider the results of the second sectional analysis performed, i.e., the one made for the "Spanish" level of the "Language" variable. Table 4.15 is the corresponding Table of Analysis of Variance.

Source	Sum of Squares	d.f.	Variance Estimate	<u>F</u>
"Spanish"	3.5999	1	3.5999	5.8544*
"Level"	3.125	2	1.5625	2.5410 (ns.)
"Spanish x Level"	.6417	2	.3208	
Residual	51.6584	84	.6149	
Total	59.025	89		

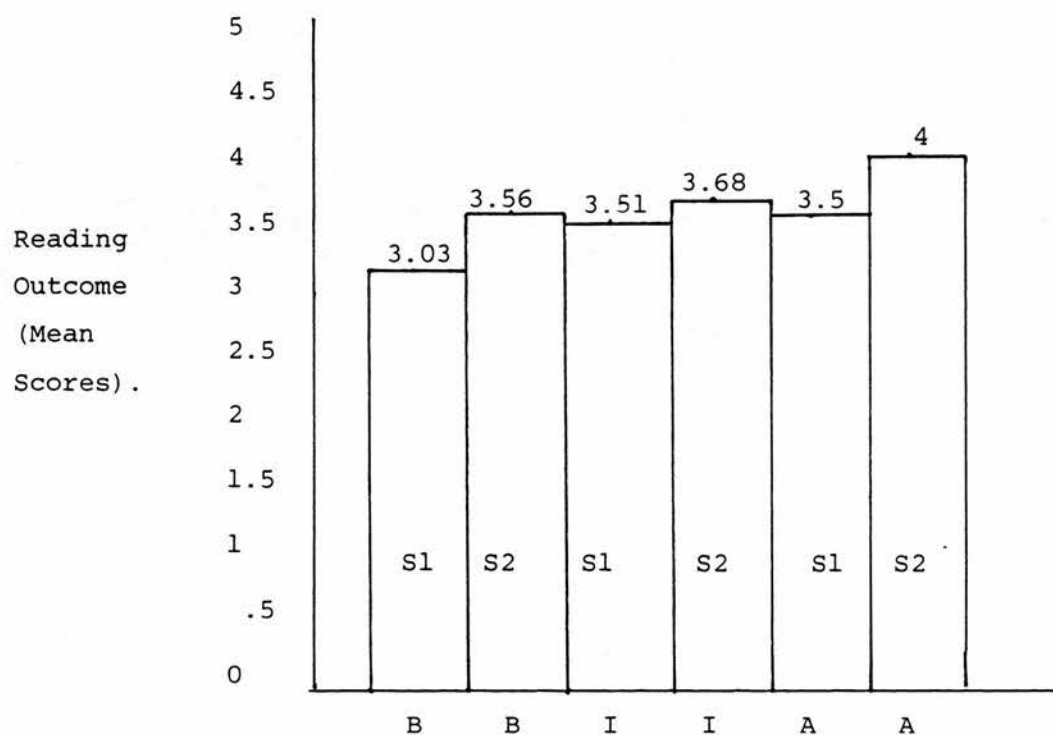
Figure 4.15 Table of Analysis of Variance.
Breakdown ANOVA by "Language:" "Spanish."

Since the variance estimate associated with the interaction "Spanish x Level" was smaller than that associated with the "Residual," the F test was not applied. The "Spanish First" and the "Spanish Second" levels of the "Spanish" variable did not interact with the "Advanced," "Intermediate" and "Beginners" levels of the "Level" (of Proficiency) variable.

On the other hand, the variance estimate associated with the main factor "Level" (of Proficiency) was greater than that associated with the "Residual." However, the difference did not prove to be statistically significant. In other words, having an "Advanced," "Intermediate" or "Beginners" level of proficiency in English did not have an effect on the subjects' performance on the Spanish test.

Finally, the variance ratio of the "Spanish" variable was significant at the 5% level, but not at the 1% level. The observed value of F was 5.8544 for 1 and 84 degrees of freedom, while the table values for the 5% and 1% levels are 4.00 and 7.08 respectively, for 1 and 60 degrees of freedom. (Lindquist, 1953,p.44). This means that having been in the condition "Spanish First" differed significantly from having been in the "Spanish Second" one.

Figure 4.22 illustrates these results. Once again, the graph is based on the means of the corresponding groups.



Where:

S1 = "Spanish First"
 S2 = "Spanish Second"
 B = "Beginners"
 I = "Intermediate"
 A = "Advanced"

Figure 4.22 Reading Outcome as Determined by the "Beginners", "Intermediate" and "Advanced" Levels of the "Level" (of Proficiency) Variable, and by the "Spanish First" and "Spanish Second" Levels of the "Spanish" Variable.

4.4.2.2 Breakdown Analysis of Variance by "Level" (of Proficiency)

Consideration will be given in this section to the three separate analyses that focus on the variable "Level" (of Proficiency). In other words, the way in which the variables "Language" (of Test) and "Order" (of Administration) affected the variable "Level" (of Proficiency) will be examined for the "Advanced," "Intermediate" and "Beginners" groups separately.

Accordingly, the first sectional analysis of the breakdown analysis by "Level" will be that one made for all the entries in Table M.1 (Appendix M) which were recorded at the level "Advanced." The second one will be made for all the entries recorded in the same table at the level "Intermediate," and the third one will be made for all the entries recorded at the level "Beginners."

Each of these analyses is again a two-factor analysis: "Language" (of Test) x "Order" (of Administration), and since both variables have two levels each ("English"/"Spanish," and "First"/"Second," respectively), the tests used this time were 2 x 2 ANOVA's.

There are now three Tables of Analysis of Variance to be considered in this section ("Advanced," "Intermediate" and "Beginners"). They are based on the information included in Tables N.1 (Advanced) through N.3 (Advanced), N.1 (Intermediate) through N.3 (Intermediate), and N.1 (Beginners) through N.3 (Beginners) in Appendix N, respectively.

With this background information in mind, we can now turn our attention to the first one of such tables of analysis of variance, the one that focuses on what happens at the "Advanced" level. (Table 4.16).

Source	Sum of Squares	d.f.	Variance Estimate	F
"Language"	.4167	1	.4167	
"Order"	.9375	1	.9375	1.7332 (n.s.)
"Language x Order"	.9375	1	.9375	1.7332 (n.s.)
Residual	30.2917	56	.5409	
Total	32.5834	59		

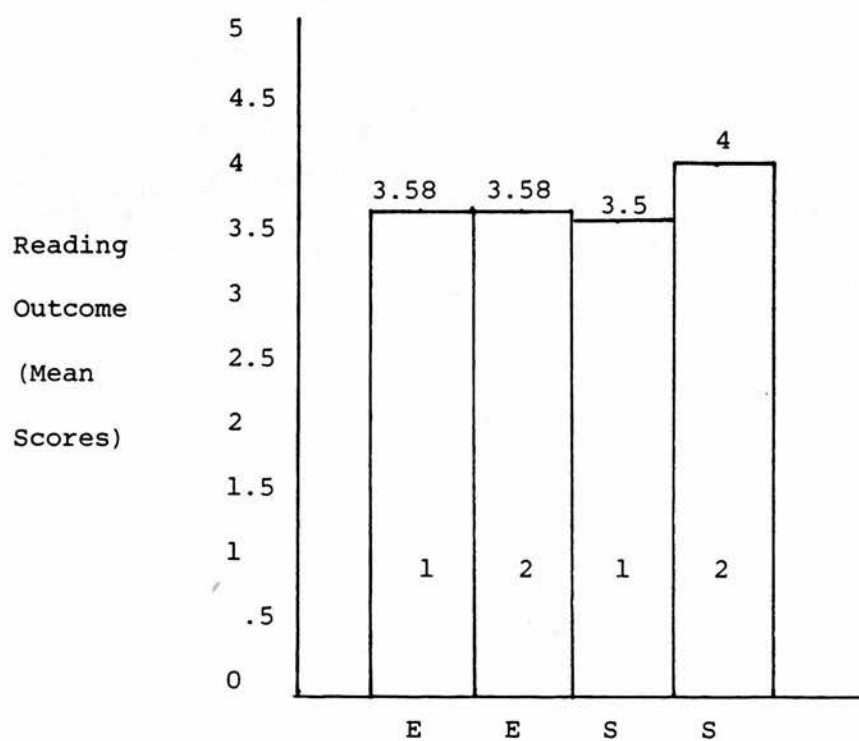
Table 4.16 Table of Analysis of Variance.
Breakdown ANOVA by "Level" (of Proficiency):
"Advanced".

It will be noticed that the F ratios worked out for the interaction "Language x Order," and for the main effect "Order" were not statistically significant.

It can also be noticed that the variance estimate associated with the main factor "Language" was smaller than that of the "Residual," and that therefore, the F ratio was not even calculated for that factor.

These results taken together indicate that at the "Advanced" level, there is no significant difference between the two languages in question, nor between the two orders of administration.

Figure 4.23 illustrates these results. The values shown are those of the means of each of the groups being contrasted.



Where:

- 1 = "First"
- 2 = "Second"
- E = "English"
- S = "Spanish"

Figure 4.23 Reading Outcome as a Function of "Language" (of Test) and "Order" (of Administration), for the "Advanced" Group.

Let us now consider what happens at the "Intermediate" level.

(Table 4.17).

Source	Sum of Squares	d.f.	Variance Estimate	<u>F</u>
"Language"	4.676	1	4.676	5.1955*
"Order"	.551	1	.551	
"Language x Order"	.0094	1	.0094	
Residual	50.40001	56	.9	
Total	55.6365	59		

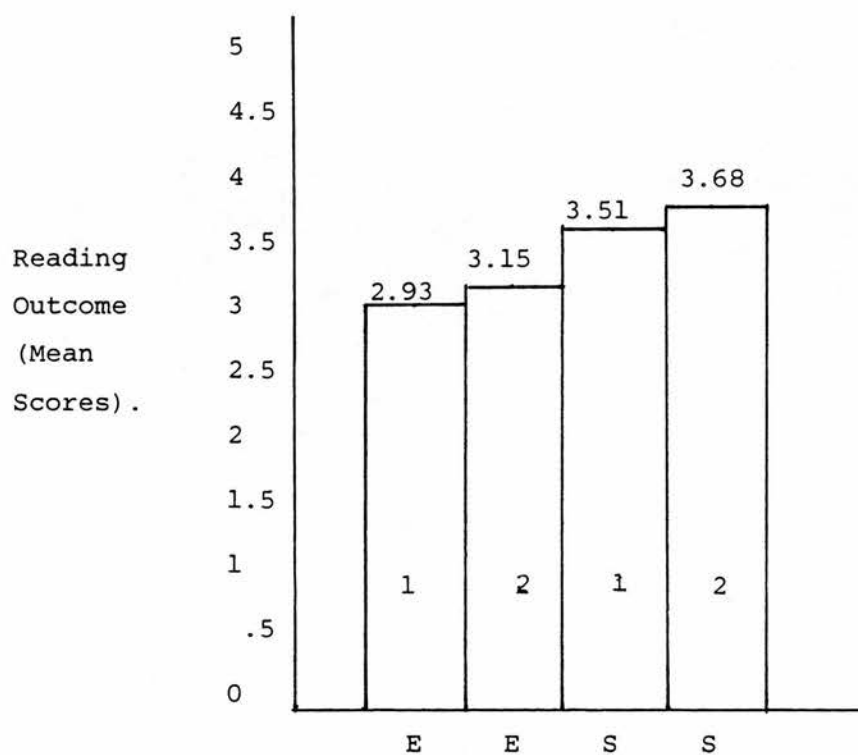
Table 4.17 Table of Analysis of Variance.
Breakdown ANOVA by "Level" (of Proficiency):
"Intermediate."

It can be observed that the F ratio was not calculated for the "Language x Order" interaction, because the variance estimate associated with it is smaller than that associated with the "Residual." As a matter of fact, the same thing happened regarding the "Order" effect; hence the blanks in the table. This means that for the "Intermediate" group neither "Order" alone, nor its interaction with "Language" affected the subjects' reading outcomes significantly.

However, the F ratio corresponding to the "Language" main effect (5.1955) for 1 and 56 degrees of freedom was greater than the table value for 1 and 40 degrees of freedom at the 5% level (4.08) (Lindquist, 1953,p.44). That is the reason behind the asterisk

in front of this \underline{F} value in the table. This indicates that for the "Intermediate" group the "Language" of the test taken does make a significant difference: their performance in the "English" test was poorer than that in the "Spanish" test.

Figure 4.24 illustrates these results, using the means of the corresponding groups to draw the graphs.



Where:

- 1 = "First"
- 2 = "Second"
- E = "English"
- S = "Spanish"

Figure 4.24 Reading Outcome as a Function of "Language" (of Test) and "Order" (of Administration), for the "Intermediate" Group.

Finally, let us consider the results for the "Beginners" level.

(Table 4.18).

Source	Sum of Squares	d.f.	Variance Estimate	<u>F</u>
"Language"	17.0067	1	17.0667	27.7868***
"Order"	5.1042	1	5.1042	8.3103**
"Language x Order"	.0375	1	.0375	
Residual	34.4	56	.6142	
Total	56.6084	59		

Table 4.18 Table of Analysis of Variance.

Breakdown ANOVA by "Level" (of Proficiency):
"Beginners" Level.

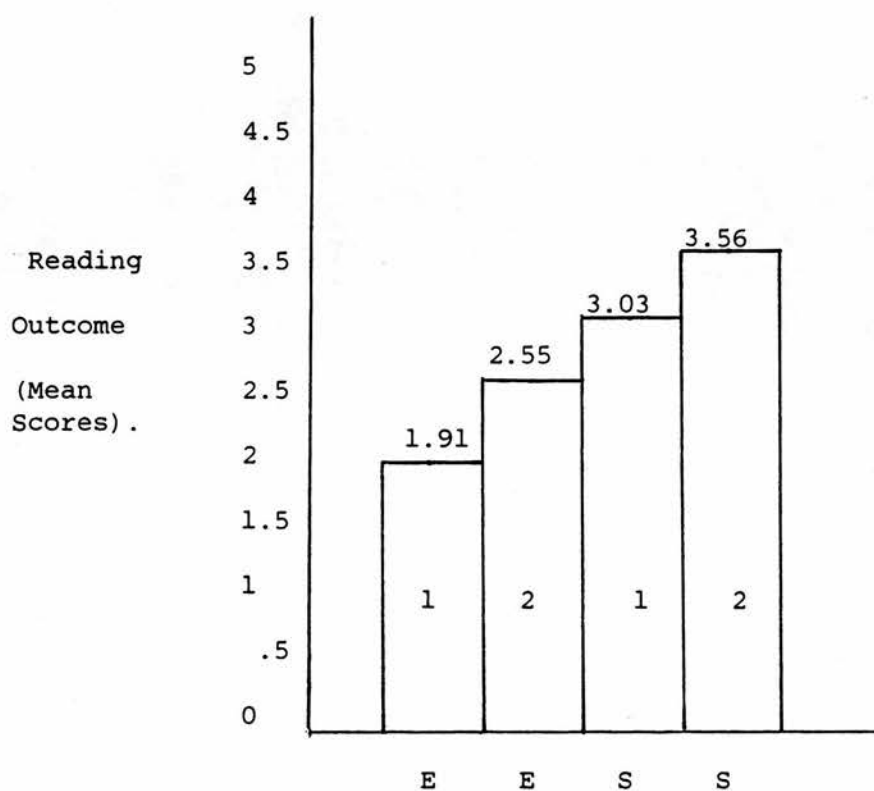
The first thing to observe in this table is the fact that the variance estimate associated with the interaction "Language x Order" was smaller than that of the "Residual," and that therefore no attempt was made to apply the F test in this occasion, hence the blank in the table.

On the other hand, both main effects were significant. In the case of the variable "Order" the F ratio of 8.3103 for 1 and 56 degrees of freedom was larger than that of the table value for 1 and 40 degrees of freedom at the 1% level: 7.31 (Lindquist, 1953,p.44); hence the two asterisks in the table.

As for the other main effect, "Language," it was significant as well. The observed F value (27.7868) for 1 and 56 degrees of freedom was much greater than the table value for 1 and 40 degrees of freedom at the 0.1% level: 12.61. (Ibid.).

These results indicate that the subjects in the "Beginners" group performed significantly better in the "Spanish" test, and in the "Second" test, regardless of whether it was the Spanish or the English test. However, the "Language" and "Order" factors did not interact, i.e., there was no differential increase.

Figure 4.25 provides an illustration of the results of this analysis. As usual, the values are the means of the groups involved.



Where:

1 = "First"

2 = "Second"

E = "English"

S = "Spanish"

Figure 4.25 Reading Outcome as a Function of "Language" (of Test) and "Order" (of Administration), for the "Beginners" Group.

A last remark regarding the results of the three analyses reported in this section. If the three tables of Analysis of Variance are compared (Tables 4.16, 4.17 and 4.18), it will be noticed that in none of them was the interaction "Language x Order" significant. It can also be observed, regarding the main effects, that an interesting pattern emerges: in the "Advanced" group neither "Order" nor "Language" reached significance; in the "Intermediate" group only "Language" did; and in the "Beginners" group both "Order" and "Language" were significant.

As mentioned before, these results will again be taken into consideration when appropriate in the following chapters.

In the next chapter, the discussion of the results of the experiment will be made from a SOLO perspective.

CHAPTER FIVEDiscussion of Results

5.1 Introduction

The discussion of the experimental results reported in the previous chapter will be made as follows. First, the results of the analyses of variance performed on the data will be discussed from a SOLO perspective. Then, consideration will be given to the main findings of the study in relation to the experimental hypotheses.

5.2 ANOVA Results and SOLO Levels

It will be remembered that when the complete ANOVA test was run on the data, it was found that neither the second order interaction ("Language x Order x Level"), nor the first order interactions "Language x Order," or "Order x Level," turned out to be significant. On the other hand, the interaction "Language x Level" was significant beyond the 0.1% level (cf. Section 4.4.1). The way in which the explanatory variables "Language" (of Test) and "Level" (of Proficiency) interacted was then analysed by means of further analyses of variance (cf. Section 4.4.2). This interaction will be discussed from a SOLO perspective in Section 5.2.1.

It will also be remembered that having found the interaction "Language x Level" significant made it invalid to test the significance of the explanatory variables ("Language of Test" and "Level of Proficiency") separately. However, since the variable

"Order of Administration" was not involved in the only interaction that reached significance, it was valid to test its significance as a main factor. Since the F test for this factor made it clear that it was significant beyond the 0.1% level (cf. Section 4.4.1), its effects on the dependent variable will also be discussed from a SOLO perspective. This will be done in Section 5.2.2.

5.2.1 Discussion of the Interaction "Language x Level" from a SOLO Perspective

The interaction "Language x Level" will be discussed in relation to the experimental hypotheses in a later section of this chapter (Section 5.4). However, it was felt that a closer look at the subjects' SOLO scores taking this interaction into account would also be profitable. The relevant information for this purpose (taken from Table 4.6) is summarized in Table 5.1.

SOLO Levels	E n g l i s h			S p a n i s h			Totals
	Begin- ners	Inter- mediate	Advan- ced	Begin- ners	Inter- mediate	Advan- ced	
1 Prestructural	2 (1.11)	1 (.55)	- -	- -	- -	- -	3 (1.66)
1.5 Transitional	3 (1.66)	- -	- -	- -	- -	- -	3 (1.66)
2 Unistructural	4 (2.22)	2 (1.11)	- -	1 (.55)	1 (.55)	- -	8 (4.44)
2.5 Transitional	11 (6.11)	6 (3.33)	3 (1.66)	2 (1.11)	4 (2.22)	2 (1.11)	28 (15.55)
3 Multistructural	5 (2.77)	4 (2.22)	1 (.55)	3 (1.66)	3 (1.66)	5 (2.77)	21 (11.66)
3.5 Transitional	1 (.55)	6 (3.33)	6 (3.33)	14 (7.77)	3 (1.66)	2 (1.11)	32 (17.77)
4 Relational	2 (1.11)	7 (3.88)	10 (5.55)	4 (2.22)	6 (3.33)	8 (4.44)	37 (20.55)
4.5 Transitional	- -	1 (.55)	6 (3.33)	3 (1.66)	6 (3.33)	6 (3.33)	22 (12.22)
5 Extended Abstract	2 (1.11)	3 (1.66)	4 (2.22)	3 (1.66)	7 (3.88)	7 (3.88)	26 (14.44)
Totals	30 (16.66)	30 (16.66)	30 (16.66)	30 (16.66)	30 (16.66)	30 (16.66)	180 (100)

Table 5.1 SOLO Scores Taking Interaction "Language x Level" into Account: Number (and Proportion) of Subjects per Category.

Table 5.1 contains the number and, in parenthesis, the proportion of subjects falling within each of the categories under consideration. It will be noticed that the effect of the variable "Order" has been left out of this table, since here the focus of attention is the interaction "Language x Level." Thus, the figures in each column correspond to all the subjects in that category, regardless of the order in which they took the test in question. That is why, for example, although 15 "Beginners" took the English test first, and 15 "Beginners" took it second, the column total for the category "English-Beginners" is 30.

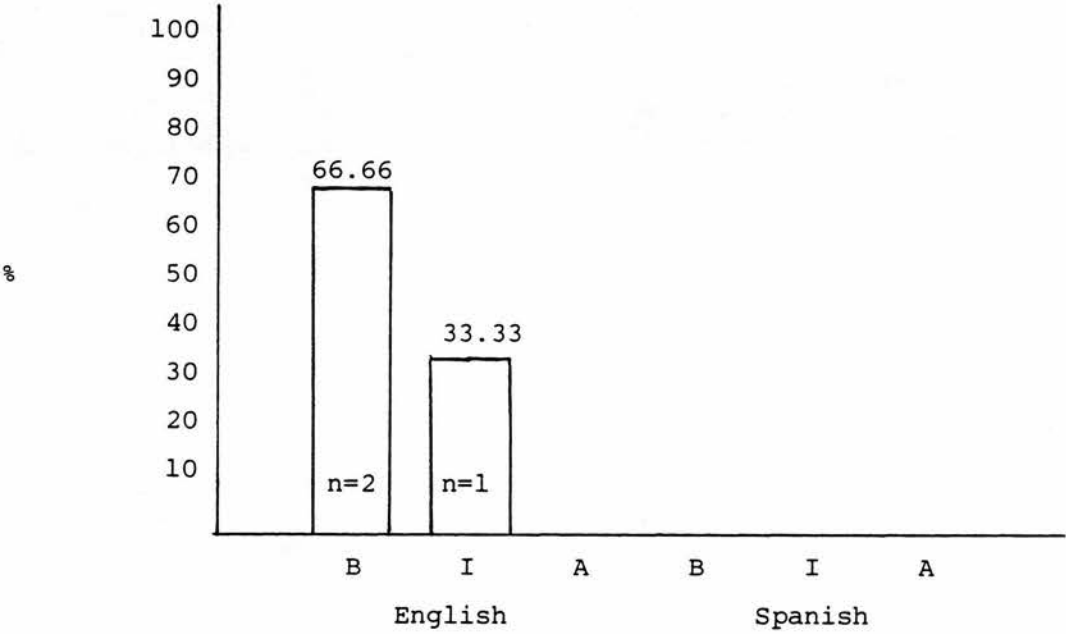
5.2.1.1 SOLO Levels

The information presented in Table 5.1 regarding the number of subjects who fell within each of the categories of the SOLO Taxonomy has been broken down to facilitate its discussion considering each level at a time. This way of organising the data is illustrated in Figures 5.1 through 5.9 (each per SOLO level).

The total number of subjects per SOLO level (N) is shown at the top of each graph, while the number inside the bars (n) corresponds to the category immediately below it. The discussion of the results observed at each SOLO level follows the corresponding graph. The transcripts of the subjects' responses referred to in the discussion can be found in Appendix O.

SOLO Level: Prestructural.

(N=3)



Where:

- B = "Beginners"
- I = "Intermediate"
- A = "Advanced"

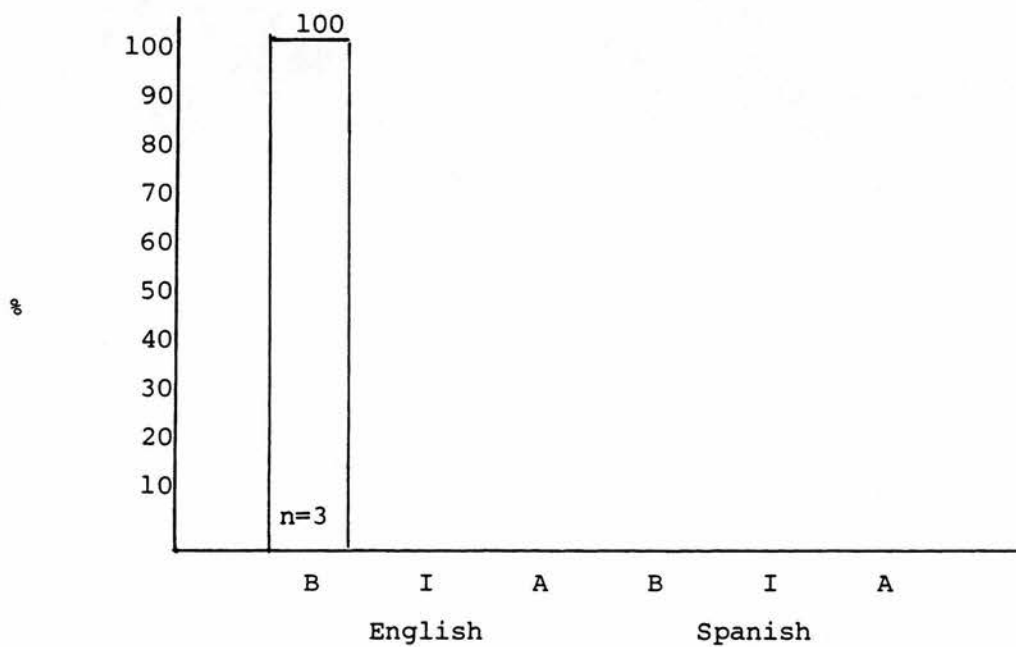
Figure 5.1 Prestructural Responses as a Function of "Language x Level".

As can be seen in Figure 5.1, there were only three subjects (out of a total of ninety) whose responses fell within the "Pre-structural" level. They all got this mark in the English test: two of them were in the "Beginners" group, and the other one was in the "Intermediate" group.

The response given by Subject 16 illustrates this response level. Her summary represents a series of attempts to get at the meaning of the original, but these attempts - as the subject herself realizes in her comments - are quite unsuccessful due to her lack of linguistic resources. In her summary she makes use - rather unsuccessfully, unfortunately - of the strategy of "borrowing" (Corder, 1981) when she realizes that the term "environmental" is quite important in the text and cannot find an adequate translation into Spanish. As already pointed out, the overall result is a series of rather unclear, unconnected remarks. The subject does not provide any comments on the content of the original, though she takes the opportunity to point out that she found the task quite difficult.

SOLO Level: Transitional (from Prestructural to
Unistructural.

(N = 3)



Where:

B = "Beginners"

I = "Intermediate"

A = "Advanced"

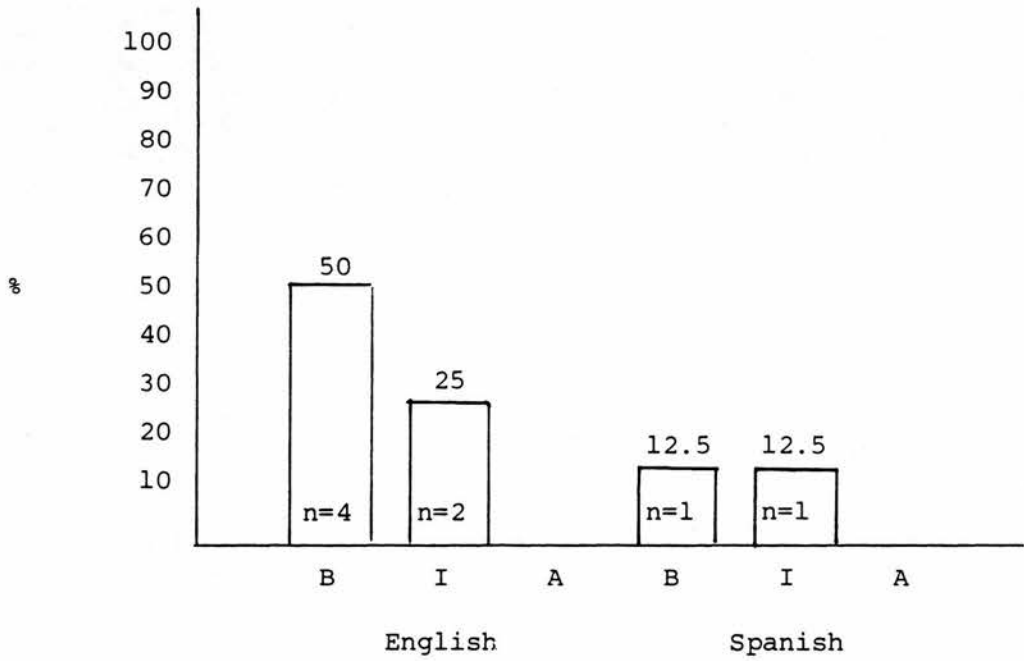
Figure 5.2 Transitional Responses (from Prestructural to Unistructural) as a Function of "Language x Level."

Figure 5.2 shows that only three responses fell into the "Transitional" category (from Prestructural to Uniststructural), and that they all were from the "Beginners" group, again in the English test.

The response given by Subject 8 illustrates this level. The summary deals with one aspect of the original basically, i.e., that of population growth. However, the fact that it deals with it only partially (the original refers to population growth in general, not only in "industrialized societies"), makes it somehow erroneous. This accounts for its not being included in the "Uniststructural" level. Again, there is no comment on the content of the original, though the fact that it was difficult for the subject is pointed out.

SOLO Level: Unistructural

(N = 8)



Where:

- B = "Beginners"
- I = "Intermediate"
- A = "Advanced"

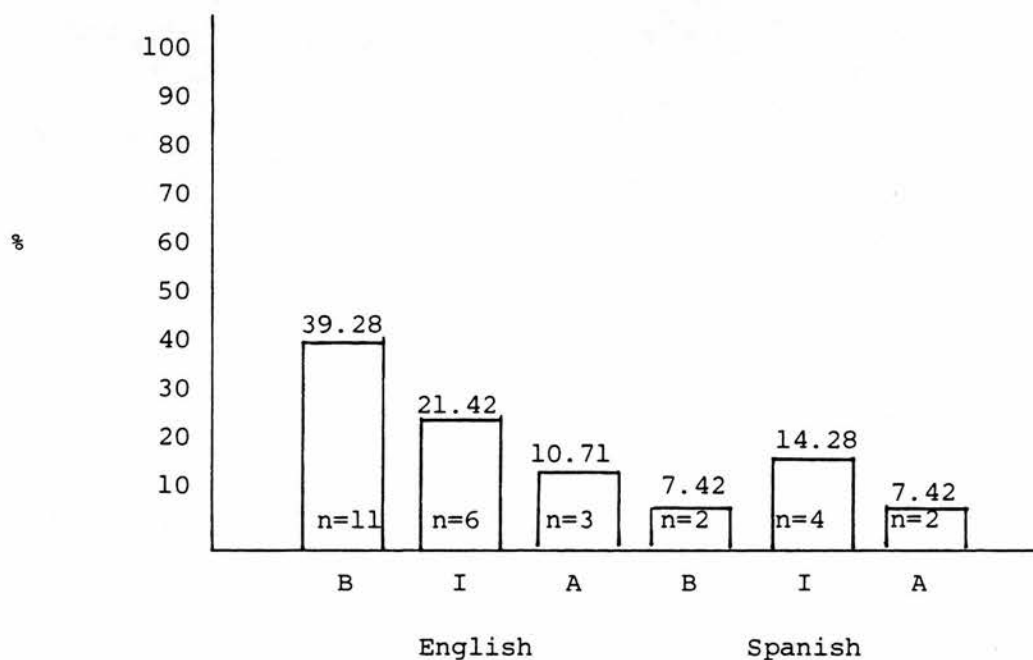
Figure 5.3 Unistructural Responses as a Function of "Language x Level."

Although some of the responses at the "Unistructural" level were given by subjects in the Spanish test, still the majority of such answers - 75% of them - were given in the English test (cf. Figure 5.3). It should be noticed that none of the subjects in the "Advanced" group got this mark.

A representative response of this level is that given in the English test by Subject 65 (from the "Beginners" group). The summary deals - correctly - with only one aspect of the problem discussed in the original text; namely, that of "economic growth." Although this time the subject includes her own comments in the corresponding section of the test, she does not go beyond what she had already dealt with in the summary. In other words, her "comments" are nothing more than a restatement of the problem she mentioned in the summary.

SOLO Level: Transitional (from Unistructural to Multistructural).

(N = 28)



Where:

B = "Beginners"

I = "Intermediate"

A = "Advanced"

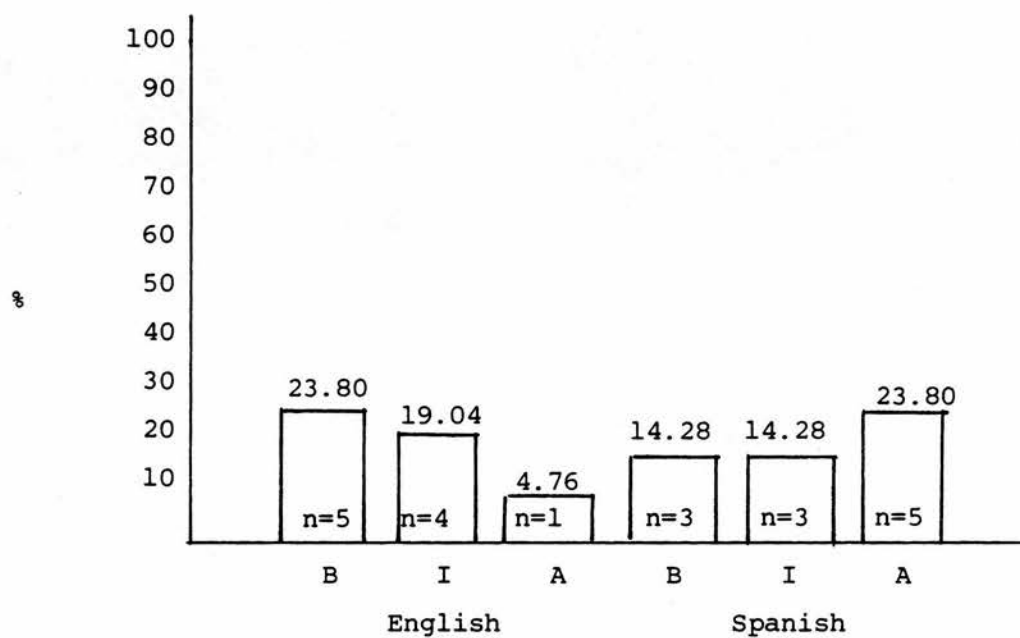
Figure 5.4 Transitional Responses (from Unistructural to Multistructural) as a Function of "Language x Level."

Although "Transitional" responses (from unistructural to multi-structural) were given by subjects in both the English and the Spanish tests, most of them (about 70%) were given by subjects in the English test, and the majority corresponded to subjects in the "Beginners" group.

The response given by Subject 70, a "Beginner" student, in the English test illustrates this level. Three aspects of the original are included in the summary: the problem of population growth, the "no-growth" approach mentioned in the text, and the danger of self-destruction. However, there are some mistakes (e.g., the author does not speak about "the tragic population growth," nor does he state that "some countries will now change to a no-growth approach"). Also, the information included is sometimes only partially true (e.g., the original text refers to population growth in general, not only in "industrialized societies"). The subject's comments on the content of the text do not go beyond the information dealt with in the summary. Once again, the opportunity is taken to mention the fact that the text was a bit difficult, but also that some "key words" were very useful to get a basic idea of its content. So, although more than one aspect is dealt with, the problems already mentioned (mistaken and partially true information) prevent this response from qualifying for the "Multistructural" level.

SOLO Level: Multistructural.

(N = 21)



Where:

B = "Beginners"

I = "Intermediate"

A = "Advanced"

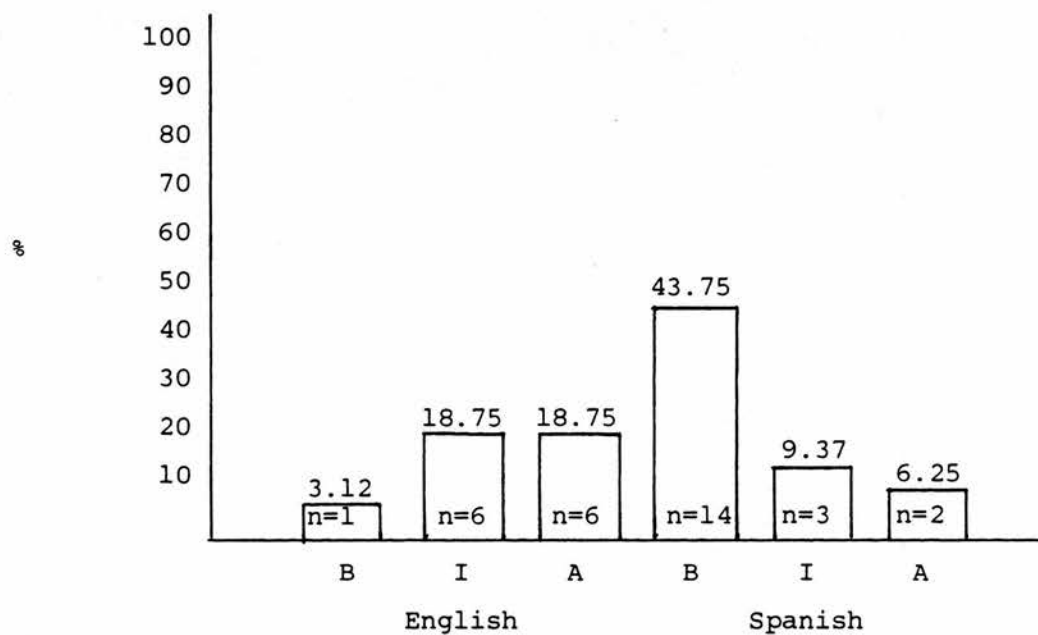
Figure 5.5 Multistructural Responses as a Function of
"Language x Level."

As far as the "Multistructural" level is concerned, a more balanced situation can be observed regarding both, the language of the test taken, and the level of proficiency of the subjects (cf. Figure 5.5).

The response given by Subject 62, from the "Intermediate" group, in the Spanish test, has been chosen as representative of this level. In her summary, she includes relevant data, but basically related to only one problem dealt with in the original, namely that of population growth. The aspect of economic growth is mentioned, though just in passing. In her comments, the subject also makes reference to the environment. However, these different parts of the original are not clearly connected to one another in the subject's response.

SOLO Level: Transitional (from Multistructural
to Relational).

(N = 32)



Where:

B = "Beginners"

I = "Intermediate"

A = "Advanced"

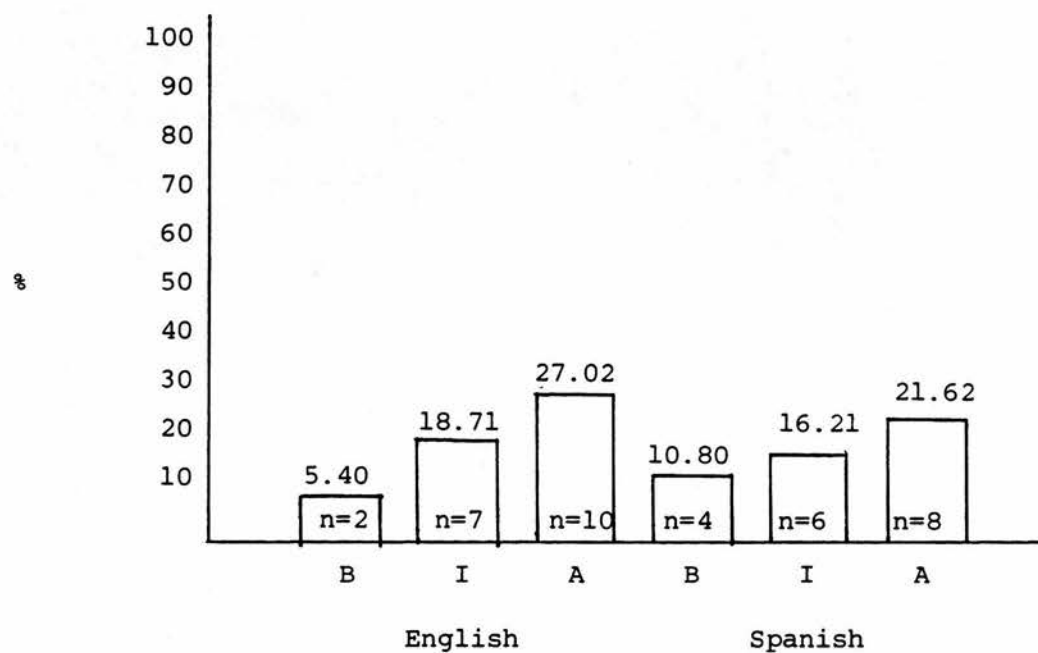
Figure 5.6 Transitional Responses (from Multistructural to Relational) as a Function of "Language x Level."

As Figure 5.6 shows, the largest group of subjects who obtained this mark (i.e., Transitional, from Multistructural to Relational) was that of "Beginners" in the Spanish test. It may also be observed, regarding the results in the English test, that while the number of "Beginners" has gone down, that of "Advanced" students has grown larger (in comparison with the preceding levels of the taxonomy).

The subject whose response has been chosen to illustrate this level (Subject 81) belongs to the "Intermediate" group, and obtained this mark in the English test. Although reference is made in his response to several relevant aspects of the original related to the problem of population growth, the subject barely mentions those aspects that have to do with the growing demand for material goods. So, although he acknowledges the need to find a balance between man and the environment, he fails to establish clearly the argument developed in the original. This prevents his answer from reaching the "Relational" level of the taxonomy.

SOLO Level: Relational.

(N = 37)



Where:

B = "Beginners"

I = "Intermediate"

A = "Advanced"

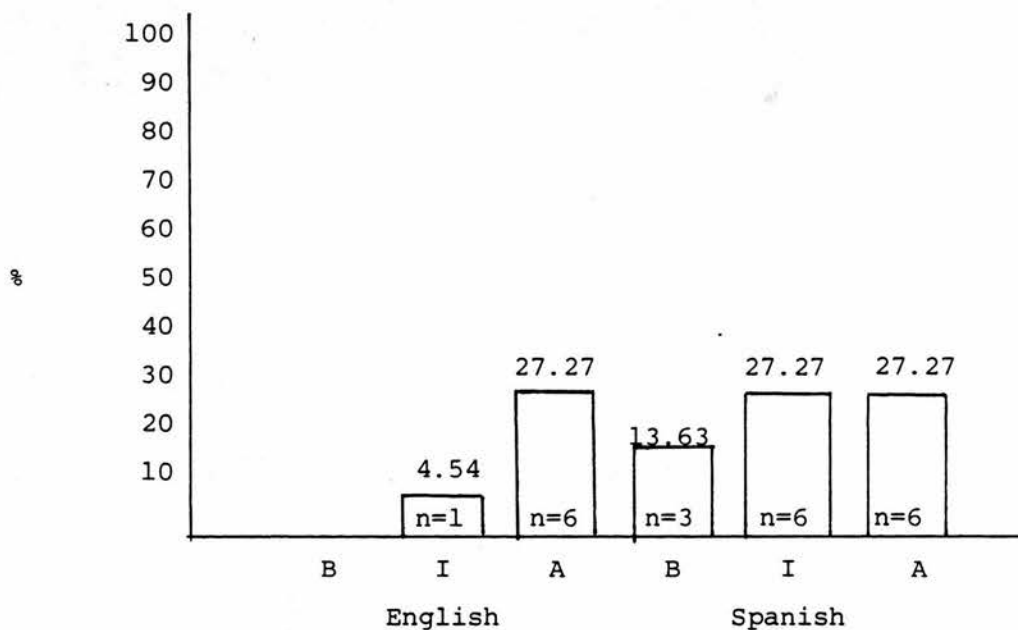
Figure 5.7 Relational Responses as a Function of "Language x Level."

The same pattern can be observed in the English and Spanish tests regarding "Relational" responses (cf. Figure 5.7). That is, in both cases there are fewer beginners than intermediate subjects, and fewer intermediate than advanced ones. It should also be pointed out that this was the modal level in this study, since more responses were awarded this mark than any other one of the taxonomy (cf. Table 5.1).

A response illustrating this level, is that given by Subject 26 (from the "Advanced" group) in the Spanish test. In her summary she establishes the relating concept required for this level; namely, that population growth as well as the growing demands for material goods must be controlled in order to achieve a balance between man and the natural environment. In her comments, the subject fails to go beyond this level.

SOLO Level: Transitional (from Relational to
Extended Abstract).

(N = 22)



Where:

B= "Beginners"

I= "Intermediate"

A= "Advanced"

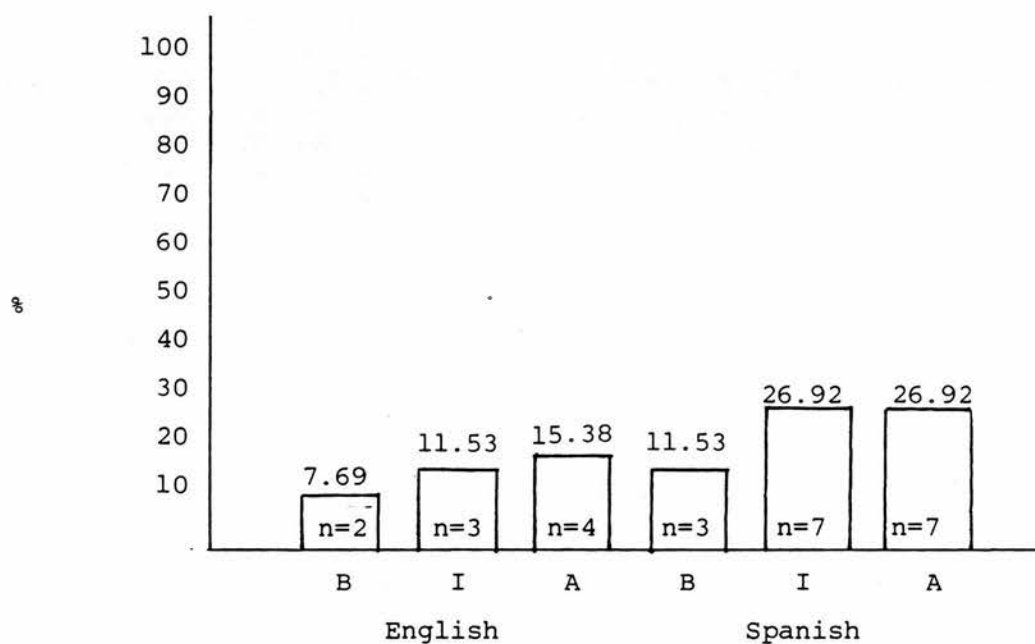
Figure 5.8 Transitional Responses (from Relational to Extended Abstract) as a Function of "Language x Level."

As far as the "Transitional" level (from "Relational" to "Extended Abstract") is concerned, it can be observed that the majority of the subjects who got it, did so in the Spanish test (approximately 68%). No beginner reached it in the English test; only one intermediate subject did, and the number of advanced students who got it was the same as that in each of the "Intermediate" and "Advanced" groups in the Spanish test.

An example of this level of response is provided by Subject 67 (an advanced student) in the English test. Having established the relating concept required in this test for a "Relational" mark in the summary, this subject makes reference in his comments to another related problem not dealt with in the original (namely, that of excessive urban growth observed mainly in the Third World). He also states that the socio-economic stance taken to deal with the ecological problem is a positive aspect of the text. However, neither of these two points contributed by the subject is further developed, thus failing to reach the Extended Abstract category.

SOLO Level: Extended Abstract.

(N = 26)



Where:

B = "Beginners"

I = "Intermediate"

A = "Advanced"

Figure 5.9 Extended Abstract Responses as a Function of
"Language x Level."

Figure 5.9 shows that the majority of subjects who reached the Extended Abstract level did so in the Spanish test (about 65%) and that the number of those who did so in the English test increased along with the subjects' level of proficiency in the language. On the other hand, the number of "Intermediate" and "Advanced" subjects who got it in the Spanish test was the same in these two groups.

On this occasion, the response given by Subject 43 (an advanced student) in the Spanish test will provide an example of the responses observed at the Extended Abstract level. In his summary, the subject establishes the relating concept required to reach the "Relational" level in this test. Then, in his comments, he states his agreement with the new approach proposed by the author in the original, but expresses his view that it will be difficult to adopt it since it implies a large scale change of the present economic state of affairs. However, he suggests as a possible first step the need to change consumers' attitudes. This change can affect the market, and thus the present production pattern. It can be seen, then, that not only has the subject understood the argument presented in the text, but has also gone a step farther by making a relevant suggestion.

5.2.1.2 Categories Resulting from the "Language of Test
x Level of Proficiency" Interaction

In this section consideration will be given to an important aspect of the results; namely, the proportion of subjects who reached the different SOLO levels within each of the categories resulting from the interaction of the explanatory variables "Language of Test" and "Level of Proficiency." These categories are: "English-Beginners," "English-Intermediate," "English-Advanced," "Spanish-Beginners," "Spanish-Intermediate" and "Spanish-Advanced." They correspond to each of the columns in Table 5.1, from which the information relevant for this discussion has been taken.

Figures 5.10 through 5.15 contain graphic representations of the results for each of the categories in question. In each of them, the total number of subjects (N) is 30, and the number of subjects per SOLO level (n) is shown inside the corresponding division. Let us now compare what happened at each proficiency level in the English and Spanish tests.

Figures 5.10 and 5.11 contain information about the way in which the group of beginners performed in the English and Spanish tests, respectively.

ENGLISH - BEGINNERS.

(N=30)

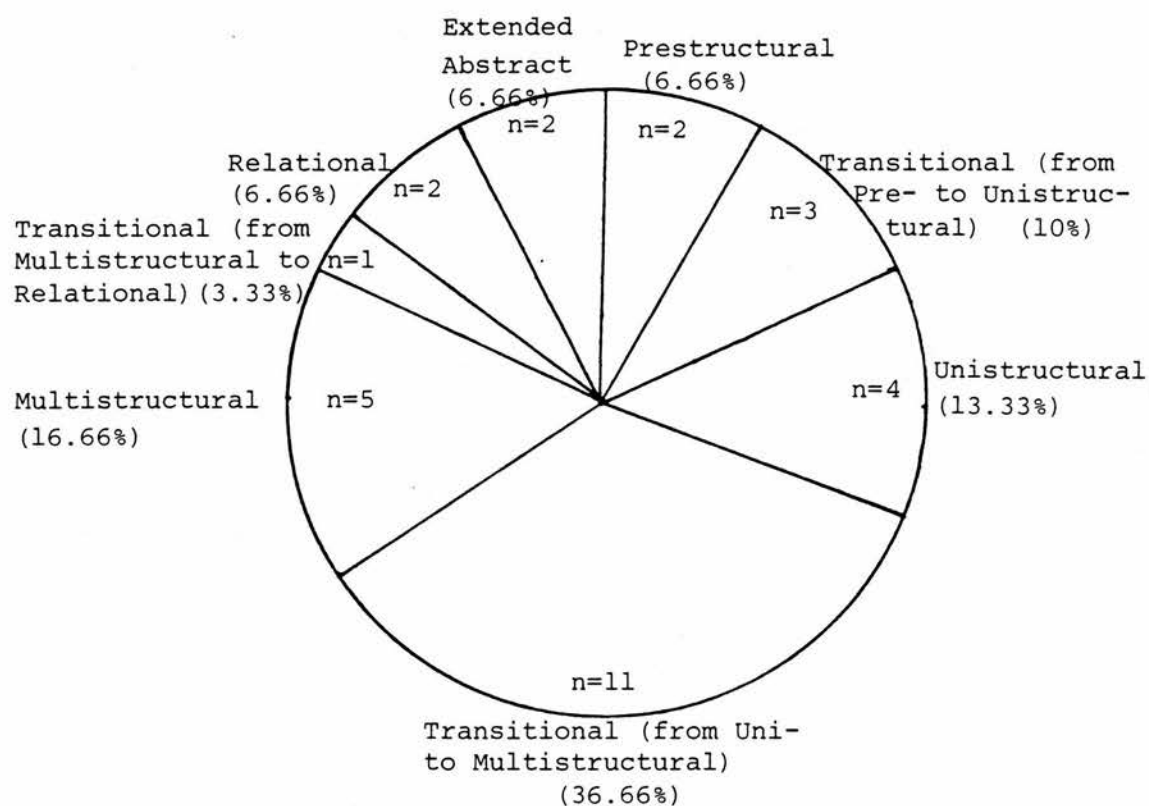


Figure 5.10 SOLO Levels as a Function of the Interaction "Language x Level:" English Test, Beginners Group.

SPANISH - BEGINNERS

(N=30)

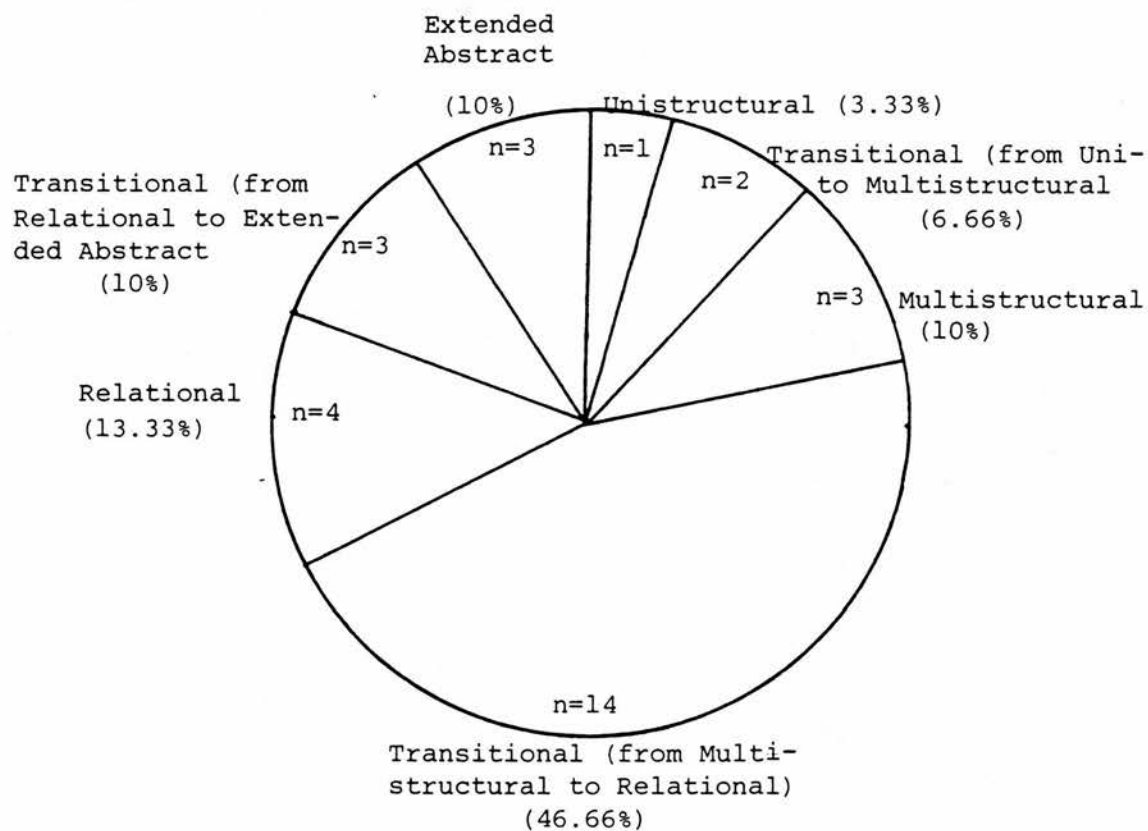


Figure 5.11 SOLO Levels as a Function of the Interaction "Language x Level:" Spanish Test, Beginners Group.

It will be noticed that while some subjects in the group of beginners obtained marks below the Unistructural level in the English test (16.66%), no such marks were recorded in the Spanish test. On the other hand, ten per cent of the subjects fell within the Transitional (from Relational to Extended Abstract) category in the Spanish test, while no subjects got this score in the English test. Finally, it should be pointed out that the median score in the English test corresponded to the Transitional (from Unistructural to Multistructural) level, while in the Spanish test it was one level higher (Transitional from Multistructural to Relational level).

ENGLISH - INTERMEDIATE

(N=30)

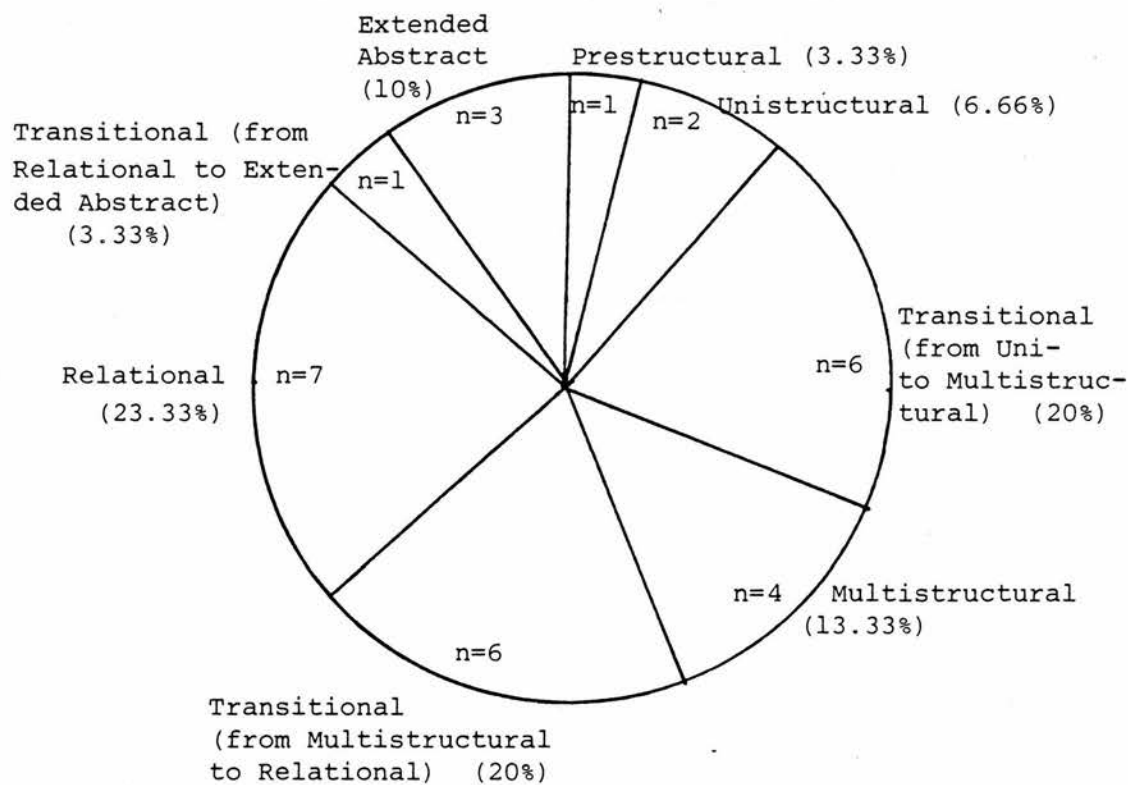


Figure 5.12 SOLO Levels as a Function of the Interaction
 "Language x Level:" English Test,
 Intermediate Group.

SPANISH - INTERMEDIATE

(N=30)

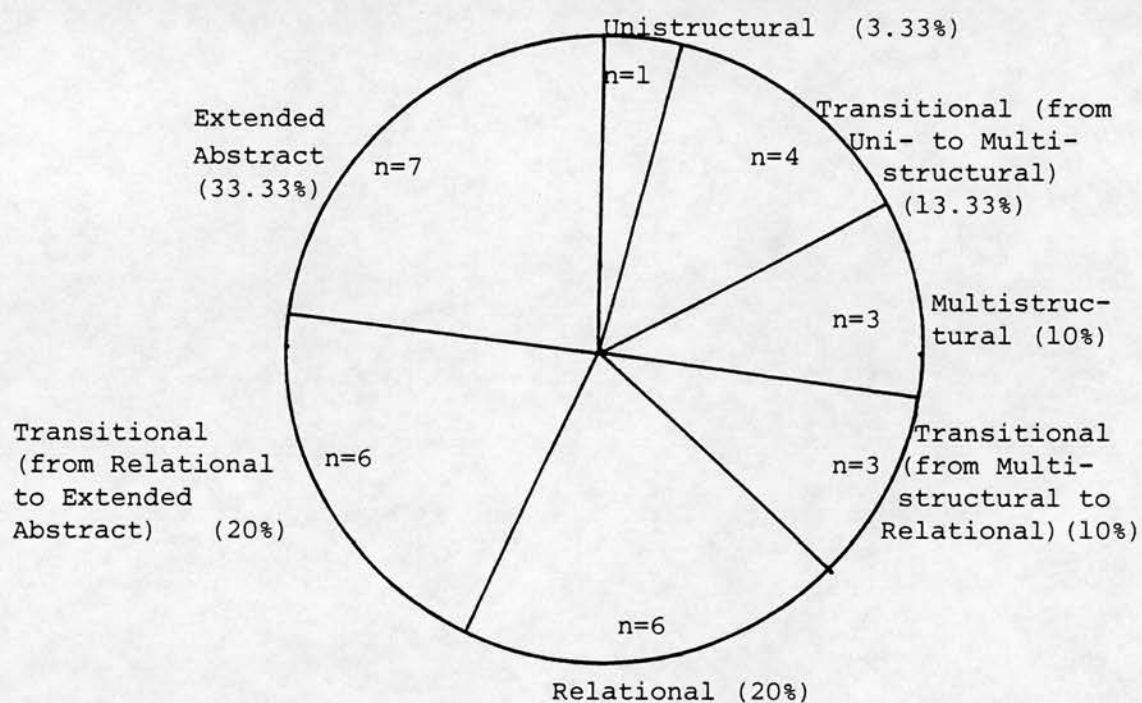


Figure 5.13 SOLO Levels as a Function of the Interaction "Language x Level:" Spanish Test, Intermediate Group.

Comparing the results of the intermediate group to those of the beginners, a general movement upwards can be observed, more subjects reaching the higher levels of the taxonomy in both tests.

However, still one subject scored below the Unistructural level in the English test (Figure 5.12), while once more no one did so in the Spanish test (Figure 5.13).

It should also be pointed out that the median went up in both tests. In the case of the English test it went one level up (Transitional from Unistructural to Multistructural → Transitional from Multistructural to Relational), whereas in the Spanish one it moved only half a level up (Transitional from Multistructural to Relational → Relational).

ENGLISH - ADVANCED

(N=30)

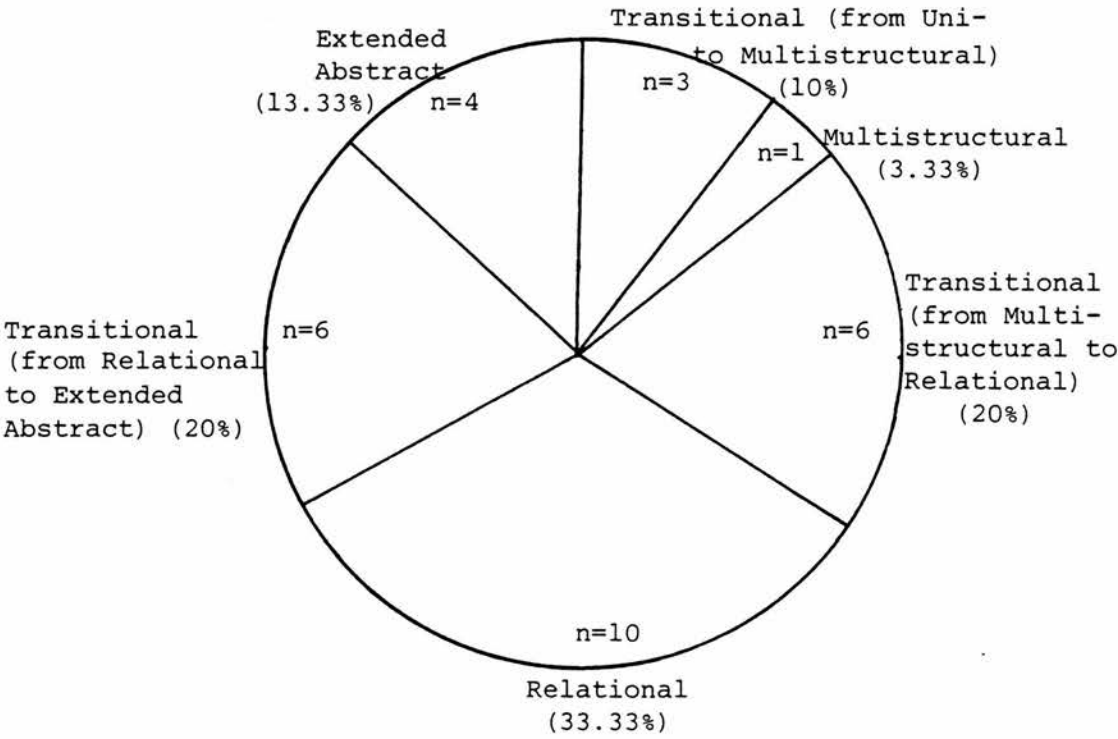


Figure 5.14 SOLO Levels as a Function of the Interaction "Language x Level:" English Test, Advanced Group.

SPANISH - ADVANCED

(N=30)

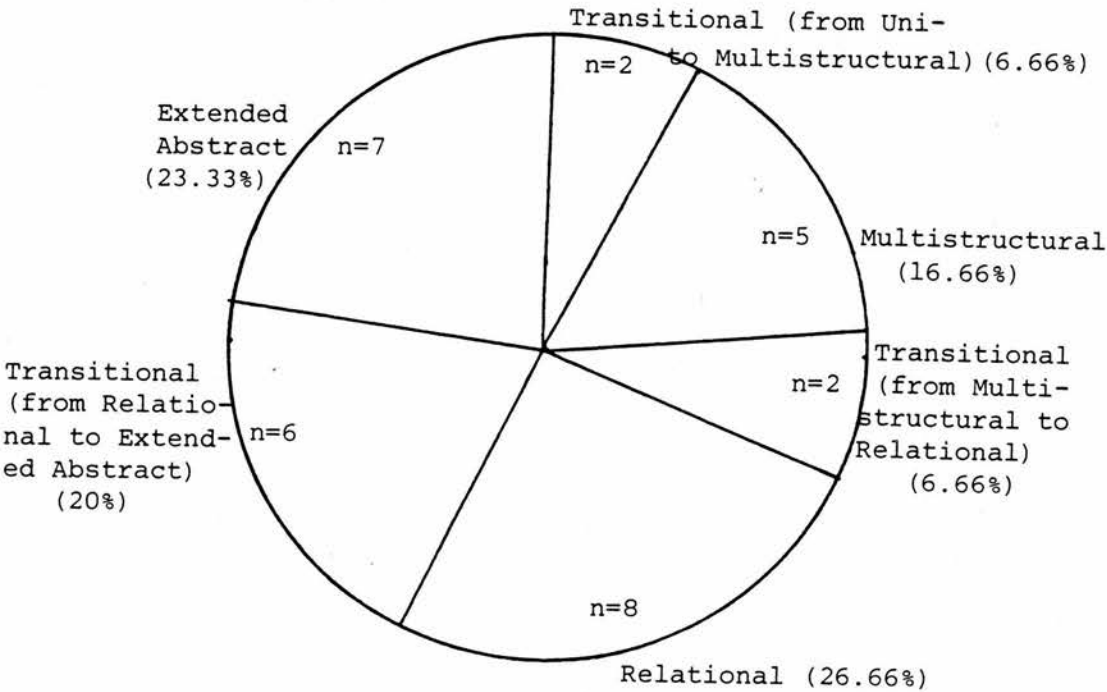


Figure 5.15 SOLO Levels as a function of the Interaction "Language x Level:" Spanish Test, Advanced Group.

Let us now look at the results obtained by the advanced subjects in the English (Figure 5.14) and Spanish (Figure 5.15) tests.

The upward trend will again be noticed, more subjects reaching higher levels in this group than in the previous ones.

It is important to notice that in the case of the advanced group there is not much difference between the results in the English and Spanish tests. In neither test, for example, did a subject score below the Transitional - from Unistructural to Multistructural - level. The median was the same in both tests (Relational), and the spread of scores was very similar.

A final remark regarding the SOLO scores obtained by the subjects in both tests refers to the fact that no "upper ceiling" was observed. That is to say, subjects from all the groups under consideration reached the Extended Abstract level. On the other hand, there was a "lower ceiling," since no subject scored below the Unistructural level in the Spanish test, and no subject in the advanced group scored below the Transitional from Unistructural to Multistructural level in either the English or the Spanish tests.

These observations lend support to the view that limited linguistic resources in the foreign language do not necessarily prevent good readers from achieving high quality standards (hence the Extended Abstract responses obtained by subjects in even the "Beginners" group in the English test).

On the other hand, having enough linguistic resources at one's disposal contributes to maintaining a minimum quality level in one's reading (hence the "lower ceiling" for all the groups in the Spanish test, and for the "Advanced" group in the English one).

5.2.2 Discussion of the "Order of Administration"

Effect from a SOLO Perspective

Having examined from a SOLO perspective the effects that the interaction "Language x Level" had on the quality of the subjects' reading outcomes, consideration will now be given to those of the main factor "Order of Administration." First, the overall variation observed in SOLO levels from the first test administered to the second will be discussed, and then attention will be paid to the way in which individual subjects performed in each test.

5.2.2.1 Overall Variation in SOLO Terms

Table 5.2 (based on the data presented in Table 4.6) contains information about the number and proportion of subjects who reached each of the different SOLO levels in the first and second test taken, regardless of the language of the test or their level of proficiency.

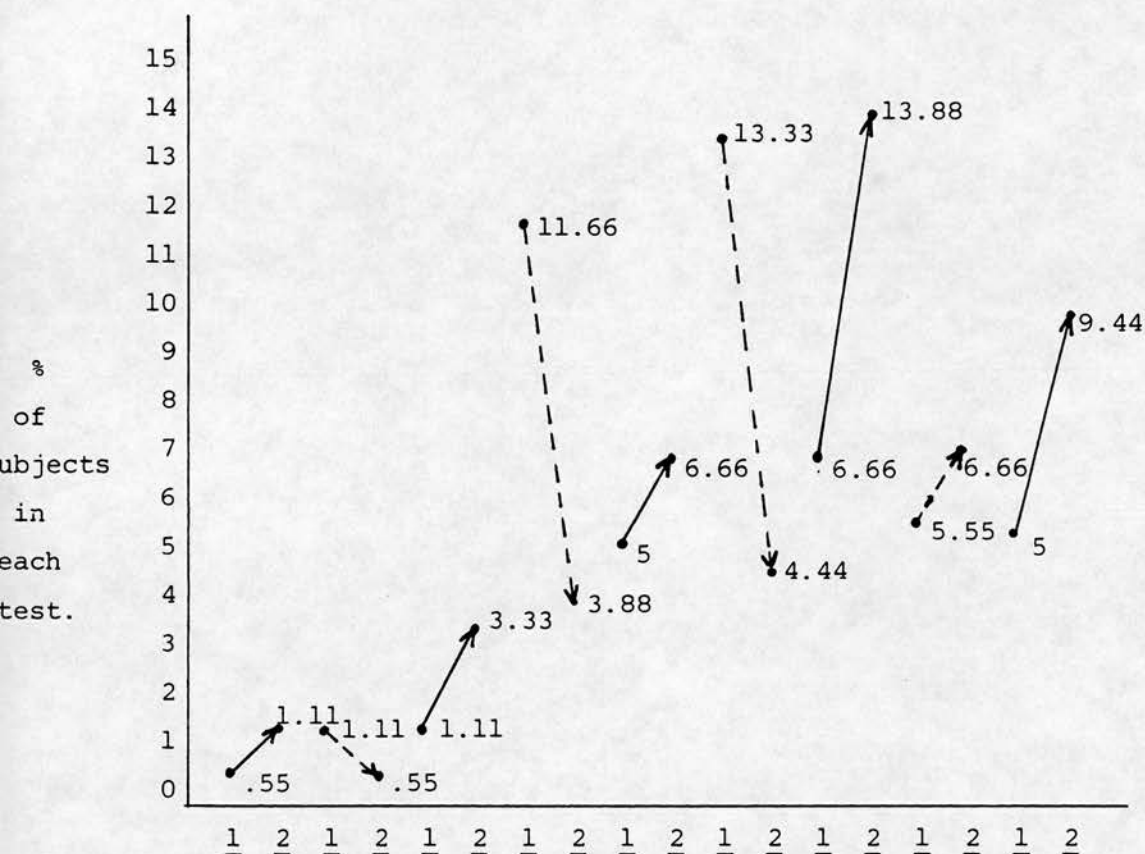
SOLO Levels	First Test	Second Test	Totals
1 Prestructural	¹ (.55)	² (1.11)	³ (1.66)
1.5 Transitional	² (1.11)	¹ (.55)	³ (1.66)
2 Unistructural	² (1.11)	⁶ (3.33)	⁸ (4.44)
2.5 Transitional	²¹ (11.66)	⁷ (3.88)	²⁸ (15.55)
3 Multistructural	⁹ (.5)	¹² (6.66)	²¹ (11.66)
3.5 Transitional	²⁴ (13.33)	⁸ (4.44)	³² (17.77)
4 Relational	¹² (6.66)	²⁵ (13.88)	³⁷ (20.55)
4.5 Transitional	¹⁰ (5.55)	¹² (6.66)	²² (12.22)
5 Extended Abstract	⁹ (5)	¹⁷ (9.44)	²⁶ (14.44)
Totals	⁹⁰ (50)	⁹⁰ (50)	¹⁸⁰ (100)

Table 5.2 SOLO Scores as a Function of "Order of Administration:" Number (and Proportion) of Subjects per Category.

Taking into consideration the information in Table 5.2, it will be noticed that the median in the first test fell within the Transitional - from Multistructural to Relational - level, whereas in the second test it was half a level above, i.e., at the Relational level. It will also be noted that the median and mode coincided in each occasion.

The upward movement from Test 1 to Test 2 that accounts for the significance reached by the "Order" factor in the ANOVA test can be better observed in Figure 5.16. This figure contains a graphic representation of the proportions of subjects who reached each of the different SOLO levels, in the first and second administration of the tests.

P	T	U	T	M	T	R	T	EA
r	r	n	r	u	r	e	r	x
e	a	i	a	l	a	l	a	b
s	n	s	n	t	n	a	n	t
t	s	t	s	i	s	t	s	e
r	i	r	i	s	i	i	i	t
u	t	u	t	t	t	o	o	r
c	i	c	i	r	i	n	n	d
t	o	t	o	u	o	a	a	a
u	n	u	n	c	n	l	l	e
r	a	r	a	t	a			c
a	l	a	l	u	l			t
l		l		r				d
				a				
				l				



here:

1 = First Test

2 = Second Test

Figure 5.16 SOLO Scores as a Function of "Order of Administration." Graphic Representation of Percentages of Subjects per SOLO Level, in First and Second Tests.

The y-axis in Figure 5.16 shows the percentage of subjects who reached the various SOLO levels in each test. So, for example, .55% of the total number of subjects in Test 1 fell within the Prestructural category, whereas 1.11% of the total number of subjects in Test 2 fell in that category; the situation was reversed in the case of the Transitional - from Prestructural to Multi-structural - level, and so on.

From what has just been said, it will be clear that the x-axis represents the main factor "Order," which takes on its two values: "First" and "Second" (1 and 2 in the graph, respectively). For the sake of facilitating the comparison being made, rather than having separate graphs for each SOLO category, a single graph has been drawn. Its long x-axis accommodates all the "First" and "Second" marks necessary for the whole taxonomy (from Prestructural to Extended Abstract).

A trend that can be noticed in this graph is that in the cases of the main categories, the proportions of subjects who scored them were always larger in the second test. On the other hand, the reverse pattern is found regarding the transitional categories. That is, the proportions of subjects who scored them were smaller in the second test, except for the case of the Transitional - from Relational to Extended Abstract - level, where a slight increase was recorded in the second test. To make this contrast clearer in the graph, the differences in proportions from Test 1 to Test 2 has been represented by continuous arrows in the case of the main categories, and by broken arrows in that of the transitional ones.

It could be said then, that besides the overall improvement in the second test, the other major trend regarding differences in SOLO levels from Test 1 to Test 2 is the overall higher proportion of transitional marks in Test 1, and the corresponding overall higher proportion of main-category marks in Test 2.

It seems to be the case, then, that as far as SOLO levels is concerned, the overall carry-over effect observed in this study took the form of less uncertainty in the way the reading task was handled in the second occasion (hence, less transitional responses in Test 2).

Having examined the overall changes in SOLO levels from Test 1 to Test 2, the discussion of variation from test to test will be approached from a different - though complementary - perspective in the next section. In it, consideration will be given to individual variation in SOLO scores from one test to the other.

5.2.2.2 Individual Variation in SOLO Terms

Each subject's performance in Tests 1 and 2 was checked (Appendix L, Column VI), and the differences - or otherwise - were recorded. The resulting information was summarized in Table 5.3.

The columns in Table 5.3 have been arranged so as to facilitate the discussion of the proportions of subjects whose SOLO scores went up, went down or remained the same from Test 1 to Test 2, taking into consideration the language of the tests (English or Spanish), and the level of proficiency of the subjects (Advanced, Intermediate or Beginners). Each of these columns will be discussed in turn.

SOLO Scores (Difference)		From Test 1 (Spanish) to Test 2 (English)			From Test 1 (English) to Test 2 (Spanish)			Totals
		Advan- ced	Inter- mediate	Begin- ners	Advan- ced	Inter- mediate	Begin- ners	
↑	n %	5 (5.55)	5 (5.55)	4 (4.44)	8 (8.88)	11 (12.22)	15 (16.66)	48 (53.33)
↓	n %	5 (5.55)	7 (7.77)	11 (12.22)	3 (3.33)	4 (4.44)	0 -	30 (33.33)
=	n %	5 (5.55)	3 (3.33)	0 -	4 (4.44)	0 -	0 -	12 (13.33)
Totals	n %	15 (16.66)	15 (16.66)	15 (16.66)	15 (16.66)	15 (16.66)	15 (16.66)	90 (100)
Where: ↑ = The subject's score went up from Test 1 to Test 2. ↓ = The subject's score went down from Test 1 to Test 2. = = The subject's score did not change from Test 1 to Test 2.								

Table 5.3 Individual Variation in SOLO Scores from Test 1 to Test 2: Number (and Proportion) of Subjects per Category.

Let us start by considering what happened regarding those subjects who took Test 1 in Spanish and Test 2 in English, beginning with the "Advanced" group.

It will be noticed that this is the only group where no differential results were recorded. The observed results coincided with those that could have been expected to happen by chance: the proportion of subjects whose scores went up (5.55% of the total number of subjects in the experiment) is the same as that of the subjects whose scores went down (5.55%), or remained the same (5.55%), in this group.

Besides providing evidence of absence of bias in the sampling procedure used, these figures reveal a wide - yet balanced - range of results in terms of the variables affecting the performance of this group of subjects. For some, there was an order of administration, or learning effect: hence their better performance in the second test in spite of its being in the foreign language. For others, there was a language effect: hence their poorer performance in the English test, in spite of having taken it second. Finally, there were other subjects whose reading performance was not affected by either the language of the tests or the order in which they took them: "good" readers were "good" readers in both tests, and "poor" readers were "poor" in both tests.

The proportional variation observed in this group reflects - it seems to me - the variety of reading outcomes that one would expect to observe in the parent population when the reader's level of proficiency in the foreign language is not an obstacle.

Let me now compare the results obtained by the subjects in the "Intermediate" group to those that have just been discussed.

It will be noticed that although there is again a wide range of results, they are not as equally distributed as in the case of the "Advanced" group. Thus, the variable "Level of Proficiency" can be said to account for the differences between these two groups.

Once again there was an order of administration, or learning effect for some subjects. As a matter of fact, the proportion of subjects whose scores went up in the second test in spite of its being in the foreign language was the same in both groups (5.55%).

However, the language effect was stronger for this group, since a larger proportion of subjects (7.77%) found their reading abilities restricted when the text was in English, even if it was the second test taken.

On the other hand, there were fewer subjects in this group (3.33%) whose reading outcome was not affected by the "Language of Test" or "Order of Administration" factors, and was therefore of the same quality in the two tests taken.

The next column in Table 5.3 contains the information about the way in which the beginners who took the Spanish test first and the English test second performed in them.

The differences between this group and the two previously discussed are striking. This again can be taken as evidence of a "Level of Proficiency" effect.

There was again some order of administration - or learning - effect, some subjects (4.44%) improving in the second test in spite of being in English. However, for the majority of subjects in this group the language effect was overriding, and their results in the English test were poorer in spite of having taken it second (12.22% of subjects fell in this category).

Finally, regarding this group, no cases of subjects with the same score in both tests were observed.

Let us now turn our attention to what happened to those subjects who took the English test first, and then the Spanish one.

In the case of the "Advanced" group, a wide range of results was again recorded, although less evenly distributed than in the case of advanced subjects taking the tests in the reverse order.

There are again some subjects in this group whose reading performance improved in the second test. However, the proportion of subjects in this category (8.88%) is larger now that the second test was the Spanish one. This means that there was an interaction of the learning (order) and language effects in this case.

There were some other subjects whose reading outcome was of a poorer quality in the second test (3.33%). That there was no learning effect for these subjects is clear. But that the language effect worked in the opposite direction of what was the general trend, and was expected, in the study, is not so easy to explain. It may be necessary to go outside the framework of the study and consider the possibility of other variables - such as lack of motivation for doing a Spanish test in their English class - accounting for these subjects performing the way they did.

Finally, it was again possible to observe in this group some subjects (4.44%) performing at the same level regardless of the language of the test or of the order in which it was taken.

Let us now move on to the next column in Table 5.3, that is, to that of the intermediate subjects who took test 1 in English and Test 2 in Spanish.

Working this time our way from the bottom row upwards, it will be noticed that no cases of subjects performing at the same level in both tests were recorded in this group, and that again some subjects (4.44%) did not behave as expected, since their scores went down in the second test in spite of being in their mother tongue.

However, for a large proportion of subjects in this group, the effect of the interaction "Language of Test" and "Order of Administration" will be noticed, since they improved their performance in the second test, which was in their mother tongue. It will be noticed, also, that the proportion of subjects who behaved this way was larger than the corresponding one in the "Advanced" group. This difference can be then attributed to the intervention of the "Level of Proficiency" variable.

The last group to be considered is that of beginners, all of whom behaved in the same way: they all performed better in the Spanish test, which they took second. "Order of Administration" may have interacted with "Language of Test," but the fact that all of the subjects in this group fell in the same category, in contrast to what happened in the case of the "Intermediate" and "Advanced" groups, makes it clear that the intervention of the variable "Level of Proficiency" played an important role, stressing the effect of the "Language of Test" variable.

Having discussed from a SOLO perspective the way in which the quality of the subjects' reading outcomes was affected by the independent variables of the study, we are in a better position - it seems to me - to discuss the findings of the study in relation to the experimental hypotheses that were formulated at its beginning. This will be done in the following section..

5.3 Main Findings in Relation to Experimental Hypotheses

Each of the experimental hypotheses formulated in Section 3.2.3 is related to one of the three independent variables of the study (cf. Section 3.2.1), and - by definition - to one of the main factors in the three-way ANOVA test used in the first stage of the quantitative analysis of the data (cf. Section 4.4.1). The relationship between experimental hypotheses, independent variables and main factors in the ANOVA test can be summarized as in Table 5.4.

EXPERIMENTAL HYPOTHESES	Independent Variables	Main Factors (Three-Way ANOVA) .
<p>Ho₁: There is no significant difference in reading outcomes between L1 and FL.</p> <p>Ha₁: There is a difference in reading outcomes between L1 and FL such that the subjects will do significantly better in the L1.</p>	Language used in the Test.	Language (of Test) .
<p>Ho₂: There is no significant difference in reading outcomes between the "Beginners," "Intermediate" and "Advanced" levels of proficiency.</p> <p>Ha₂: There is a significant difference in reading outcomes in the foreign language between the "Beginners," "Intermediate" and "Advanced" levels of proficiency, such that the higher the subjects' level of proficiency in English, the better they will perform in the English test.</p>	Level of proficiency in the foreign language.	Level (of Proficiency) .
<p>Ho₃: There is no significant difference in reading outcomes between the first or second test taken.</p> <p>Ha₃: There is a difference in reading outcomes between the first and second test taken, such that the subjects will perform better in the second test they take.</p>	Order of administration of the Tests.	Order (of administration) .

Table 5.4 Relationship Between Experimental Hypotheses, Independent Variables and Main Factors in Three-way ANOVA test.

In order to discuss the main findings of the study in relation to the experimental hypotheses, a quite straight-forward procedure would be to simply consider each of the rows in Table 5.4 in turn. However, this would be inadequate, since it will be remembered that the three-way ANOVA test used permitted the exploration of interactions between the factors, thus producing a less simple set of results. So, although this table serves the purpose of making the basic relationships between hypotheses and test explicit, the discussion that follows - rather than being directly guided by the table - will be structured in accordance with the actual results of the ANOVA test.

It may be noted, though, that if the null hypotheses were submitted to test just in terms of the main factors directly associated with them, they would indeed be rejected beyond the 0.1% significance level (cf. the F ratios that correspond to the main factors in Table 4.13). Instead, however, let us turn our attention to the results of the three-way ANOVA test in question (Section 4.4.1), beginning with the interactions.

It will be remembered that neither the second order interaction "Language x Order x Level," nor the first order interactions "Language x Order" or "Order x Level" reached statistical significance, whereas the interaction "Language x Level" was significant beyond the 0.1% level.

As a result, it was appropriate to test for the significance of the main factor "Order," but not for that of the factors "Language" or "Level," separately. The interaction of these two, on the other hand, was further explored by running the corresponding breakdown analyses of variance (Section 4.4.2).

In the discussion that follows we shall consider, then, what was found out regarding the main factor "Order," and the interaction "Language x Level," in relation to the experimental hypotheses.

5.3.1 Order of Administration

The statistical significance of the F ratio corresponding to this main factor in the three-way ANOVA was found to be beyond the 0.1% level (see Table 4.13). In terms of the third experimental hypothesis, which is the one associated with this main factor (see Table 5.4), this means that the null hypothesis of no difference can be rejected. In other words, the alternative hypothesis is accepted since the subjects performed significantly better in the second test they took.

However, it must be remembered that the inclusion of the variable "Order of Administration of Tests" in the experiment stemmed from practical considerations regarding the research design chosen (cf. Sections 2.3.4, and 3.2.1). In other words, there was no other interest in the order of administration of the tests than to eliminate its effects from the "language of test," "level of proficiency" and "error" variances. As a matter of fact, this was indeed achieved. Of course, the situation would have been different if it had interacted significantly with any of the other variables; but since it did not, it was possible to separate its effect from that of the other variables involved. On the other hand, the mere fact of being able to estimate the "order of administration" variance, made it possible to separate it from the "error" variance.

Having found this main factor significant means that there was an overall improvement in the subjects' performance in the second test taken regardless of the language in which the test was written or the subjects' level of proficiency. In the context of the present discussion, this can only be taken as evidence of there having been an overall learning effect. It may also be pointed out at this stage, that from a SOLO perspective this learning effect was revealed by the smaller proportion of transitional responses in the second test (Section 5.2.2.1).

Two concluding remarks concerning this learning effect are still in order. In the first place, it must be emphasized that it was possible to isolate, and understand it, thanks to the fact that the variable "Order of Administration of the Tests" was built into the experimental design. Also, it must be remembered that statistical significance and theoretical significance are not the same thing. In other words, the statistical significance of this main effect was important due to the considerations that have just been made regarding experimental research design. However, since it did not interact with the other independent variables - the ones that were important from a theoretical perspective -, its having been statistically significant is of no major consequence from an applied linguistics perspective.

5.3.2 "Language of Test x Level of Proficiency" Interaction

We can start this discussion about the interaction "Language of Test x Level of Proficiency" by considering what it implies in terms of the experimental hypotheses.

If the individual variables "Language of Test" and "Level of Proficiency" are respectively related to the first two experimental hypotheses (see Table 5.4), the following modifications to the hypotheses can be said to arise from their interaction.

The first and second null hypotheses can be combined as follows:

Ho: There is no significant difference in reading outcomes between L1 and FL, regardless of the subjects' level of proficiency in the FL.

The corresponding alternative hypothesis will be:

Ha: There is a difference in reading outcomes between L1 and FL such that the subjects will perform better in the L1, the difference being smaller as the subjects' level of proficiency increases.

To reject, or otherwise, this null hypothesis, it will now be necessary to refer to the relevant results of the breakdown analyses of variance performed on the data (Section 4.4.2).

Let us start by considering the results of the breakdown analysis of variance that was performed focusing on the level "English" of the variable "Language of Test." (Section 4.4.2.1). It can be pointed out that the only F ratio that reached significance in this test (Table 4.14) was that of the main effect "Level" (of Proficiency), and that it did so beyond the 0.1% level. This means, of course, - as the direction of the alternative hypothesis made it clear that was expected - that the subjects' level of proficiency in the foreign language does have an effect on the quality of their reading performance in it, such that the higher the subjects' level of proficiency, the higher the quality of their reading outcomes.

Now, regarding the results of the breakdown analysis of variance focusing on the level "Spanish" of the variable "Language of Test" (Section 4.4.2.1), it will be observed - again, as one would expect - that the factor "Level" (of Proficiency in the foreign language) did not affect the subjects' performance in the Spanish test (Table 4.15).

Let us now examine the results of the breakdown analyses of variance that focused on the three levels ("Advanced," "Intermediate" and "Beginners") of the variable "Level" (of Proficiency). (Section 4.2.2.2).

Let us start with the "Advanced" level. It will be noticed (Table 4.16) that for the advanced subjects, the factor "Language" did not reach statistical significance. In other words, these subjects' reading quality was of the same standard in both English and Spanish.

On the other hand, at the "Intermediate" level, it was found that the F ratio for the main factor "Language" was significant at the 5% level (Table 4.17). That is, for the group of intermediate subjects, the language of the test taken did make a difference regarding the quality of their reading performance, such that they did better in their mother tongue.

Finally, for the group of beginners, the effect of the "Language" factor was greater than for the "Intermediate" group, since this time it reached significance at the 1% level (Table 4.18). Thus, the difference in performance between the English and Spanish tests was greater for the "Beginners" group than that observed for the "Intermediate" one.

The preceding discussion makes it clear that the null hypothesis postulated at the beginning of this section can indeed be rejected on the basis of the results of the corresponding breakdown analyses of variance.

Accordingly, the alternative hypothesis can be accepted, since there was indeed a significant difference in reading outcomes between L1 and FL, such that the subjects in the "Beginners" group performed better in their L1, while this difference decreased as the subjects' level of proficiency in the FL increased.

It should be mentioned at this stage that the way in which the interaction of the variables "Language of Test" and "Level of Proficiency" was made manifest in the subjects' reading outcomes was discussed in detail earlier in this chapter from a SOLO perspective (cf. Sections 5.2.1.2, and 5.2.2.1).

Finally, it should be pointed out that the fact that on this occasion the interaction of these variables was found significant is important not only from a statistical point of view, but more from the theoretical perspective within which the experiment originated.

CHAPTER SIX Conclusions

Chapter Six Conclusions

6.1 Introduction

This final chapter has been structured as follows. The initial sections focus on the related issues of the reliability and validity of the testing instruments used to assess the quality of the subjects' reading outcomes in this study (Sections 6.2 and 6.3, respectively). Then, the main findings of the study are summarized (Section 6.4), and finally, its possible implications are considered (Section 6.5).

6.2 Reliability

The reliability of the marking scheme used in the experiment will be the first issue to be dealt with in this concluding chapter, since this is the first requirement a testing instrument must satisfy (cf., for example, Davies, 1977b, p.57; Oller, 1979, p. 4; and, Savignon, 1983, p. 236).

Of the procedures available for determining test reliability, the one chosen for being particularly appropriate when - as in this study - essay-type items are used, was interrater reliability (cf., for example, Biggs and Collis, 1982, pp. 186-9; and, Martuza, 1977, pp. 133-8).

The degree of agreement between the two independent judges who marked the subjects' scripts was examined in both, the pilot study and the main experiment (Sections 3.3.2.5 and 4.3.2, respectively). In both cases, the correlation observed was substantial

(the respective coefficients being .7806 and .6879), and significant beyond the 0.1% level. Having found a significant correlation coefficient in the pilot study, it was decided to use the same marking scheme without modifications in the main experiment. In turn, the fact that the correlation between the two judges in the main experiment was significant beyond the 0.1% level, was taken as evidence of a high degree of interjudge reliability of the marking scheme used. These results are in agreement with Biggs and Collis's findings (1982, pp.188-9) regarding the reliability of the SOLO Taxonomy for rating essay-type items in other content areas as well.

6.3 Validity

Having provided evidence of the marking scheme reliability, it is now necessary to examine various aspects of the experiment that have to do with its validity.

First, consideration will be given to the aspects of task and test selection (Sections 6.3.1 and 6.3.2), in relation to the notions of ecological validity, i.e., of getting the experiment as close as possible to the natural setting (cf., Entwistle, 1981, p.24), and of content validity, which "is related to the question of whether the test requires the examinee to perform tasks that are really the same as or fundamentally similar to the sorts of tasks one normally performs in exhibiting the skill or ability that the test purports to measure." (Oller, 1979, p.50).

Then, the adequacy of the research instruments used for analysing the relationship between level of proficiency and quality of reading outcomes will be considered (Section 6.3.3). The issue here, then, will be the notion of construct validity, since what will be examined is the suitability of the research tools to test experimentally the hypotheses set up (cf., Cronbach, 1961, p. 106, cited in Davies, 1977b, p. 59), or, in other words, to produce meaningful results in terms of the underlying theory (cf., Davies, 1977b, pp. 62-3; and, Savignon, 1983, p. 236-8).

6.3.1 Adequacy of Task Selection

Oller (1979, p. 415) has pointed out that "if we are interested in valid tests of how people actually use language, the ultimate validity criterion is how they actually use language." It follows, it seems to me, that the first requirement a test must satisfy in this respect is to contain tasks that are realistic for the subjects taking it.

It has already been pointed out, that the tasks the subjects were required to perform in the reading comprehension tests (i.e., to state in their own words the topic of a passage, and to make comments on the material read) are tasks they can reasonably be expected to perform in the course of their university studies (Section 3.3.2.3). On the other hand, other aspects taken into account at the time of selecting these tasks were the fact that asking subjects to summarize texts has been considered more realistic than asking them to answer a set of predetermined

comprehension questions (cf., for example, Fawcett, 1979, p. 214), and the fact that they come closer to the type of questions specialists appear to prefer, as opposed to those asked by non-specialists in the subject (Zuck and Zuck, 1984, pp. 134-5). (See Section 3.3.2.3).

These arguments support the view that the tasks in question reflect the way in which people actually use language; or, more specifically, of the way in which the population sampled may be expected to use written materials during their studies. Moreover, no evidence to the contrary was found in the results of the experiment. That is, when poor performance was observed, it could generally be explained in terms of the experimental variables, and not in terms of lack of familiarity with the experimental tasks.

6.3.2 Adequacy of Text Selection

It should be pointed out that besides requiring the subjects to perform tasks which are realistic for them, it is also necessary to ensure the adequacy of the specific materials used.

In the preparation of the experiment, the following precautions were taken in this respect. First, it was made sure that the texts selected satisfied a number of selection criteria, such as being accessible to readers with no technical knowledge of the subject, and at the same time being demanding and interesting enough for subjects with a wide range of academic backgrounds. Also, it was necessary that the texts could be read as complete

and independent units, and that the subjects could perform the tasks set within the time at their disposal. (Section 3.3.2.1).

Then, a readability study was carried out in order to ensure - as much as possible - the equivalence of the texts in the English and Spanish tests. (Section 3.3.2.2).

Having decided that the texts could be taken as equivalent in terms of degree of difficulty, they were analysed from a SOLO perspective (Section 3.3.2.3), in order to produce a list of "components or dimensions...[to] enable the distinctions between the various SOLO levels to be made" (Biggs and Collis, 1982, p.165), and then, a team of specialists summarized the texts before asking the subjects to do so (Section 3.3.2.4). This was done to ensure that from the experts' point of view, it would not be unreasonable to require the subjects to perform the expected tasks with this specific set of materials (*ibid.*, p.170), thus enhancing the content validity of the tests, for as Davies puts it (1977b, pp.61-2), this can only be established "by an expert appraisal of the test content."

To explore the viability of the materials, a pilot study was subsequently carried out (Section 3.3.2.5). No reason was found thereby to suspect they would not be adequate for the experimental subjects. As a matter of fact, the results of the experiment itself supported this view. On the one hand, the tests proved to be viable. On the other, a large proportion of subjects (35 out of 90 in Test 1, and 34 out of 90 in Test 2) included explicit positive remarks about the texts at the end of their

answer sheets (qualifying them - in general - as "interesting," "useful," "relevant" or "important"). In contrast, only two subjects expressed a negative view of Text 1 (for being "rather long," and "superficial," respectively), and one complained of Text 2 for not being clearly written (in spite of having taken Test 2 in Spanish).

6.3.3 Adequacy of Research Tools in Relation to Purposes of the Study

The adequacy of the research tools used, given the purposes of the study, will be examined from three different perspectives. First, their adequacy will be considered in relation to the qualitative/quantitative aspects of the study. Then, the adequacy of their use in relation to norm-referenced and criterion-referenced evaluation will be discussed. Finally, their usefulness for analysing the data in relation to the hypothesized results will be taken into account.

6.3.3.1 Qualitative and Quantitative Analyses

The way in which the terms quantitative and qualitative were to be understood in this study was pointed out in Chapter 1 (Section 1.4). It was mentioned that these terms have been used in the educational context to refer to "how much" has been learned and "how well" it has been learned (see Biggs and Collis, 1982, p.3), and it was made clear that for the purposes of this

study "qualitative" would also be used in the sense of "how well," but that "quantitative" would instead be used to refer to "how many." In other words, it was pointed out that the quality of the subjects' reading outcomes would be analysed first, and that afterwards, the results would be submitted to a quantitative analysis to examine how many subjects in each of the contrasting groups of the experiment had behaved as hypothesized and how many had not.

The research tools chosen for these purposes of the study were the SOLO Taxonomy, and the ANOVA test. It was felt that the former would offer a suitable basic framework for assessing the quality of the subjects' reading outcomes, and that it would also provide a scale that could be used in the quantitative analysis of the data. The latter, on the other hand, was chosen due to the characteristics of the research design chosen (cf., Sections 3.2.2 and 3.2.3).

It seems to me that in practice both analytical tools indeed proved to be suitable for the purposes of the study. This view can be supported, as far as the SOLO Taxonomy is concerned, by the results (and their subsequent discussion) of the exploratory study reported in Chapter 2 (Section 2.3), and of the pilot application of the experimental materials (Section 3.3.2.5); and as far as the combined use of the SOLO Taxonomy and the ANOVA test is concerned, by those of the main experiment itself (Chapters 4 and 5).

As a matter of fact, it is my contention that owing to the way the experiment was set up, the combined use of these research

tools made it possible to go beyond the original intentions of analysing "how well" the reading tasks had been performed by "how many" subjects, for it was also possible to explore "under what circumstances" this had taken place. In other words, it was possible to determine how many subjects reached the different SOLO levels, as a function of the experimental variables "Language of Test," "Level of Proficiency" and "Order of Administration." This was basically achieved by examining the results of the various analyses of variance from a SOLO perspective, as reported in Chapter 5 (Section 5.2).

To sum up, it seems to me that the research tools chosen were adequate in relation to this particular dimension (qualitative/quantitative analyses) of the study.

6.3.3.2 Norm-Referenced and Criterion-Referenced Evaluation

In this section consideration will be given to the adequacy of the research tools used in this study in relation to norm-referenced and criterion-referenced evaluation.

In line with the convenience noted by Farr, Carey and Tone (1986, p.184) of exploring the potential and limitations of criterion-referenced measures in reading assessment, the present study explored the usefulness of the SOLO Taxonomy - which, as mentioned earlier (Section 2.1), is defined by its authors as a "criterion referenced measure of the quality of learning" (Biggs and Collis, 1982, p.8) - to assess the quality of reading outcomes.

Ingram (1977, pp.31-2) points out that in criterion-referenced evaluation,

we assume that there are two kinds of individuals, those who can carry out an operation and those who cannot. And what we are interested in is dividing them up into two entirely separate, discontinuous groups. However, we must also accept that there is a third group of people within that population; those who are currently engaged in the process of learning to carry out the operation. And in this group there will be a continuous rather than a yes/no distribution.

In this respect, the present study provides empirical evidence of the usefulness of the SOLO Taxonomy as a criterion-referenced measure of the quality of reading outcomes, since it had the dichotomizing power necessary to discriminate between those subjects who could perform the required reading tasks adequately, and those who could not.

Let us examine how this was possible regarding the specific tasks set to the subjects in this study (i.e., summarizing a text, and commenting on the material read). In the case of the summary, the dichotomy was established at the relational level, thus discriminating between those subjects who recovered the relating concept of the original text, and those who did not. And, regarding the subjects' comments, the dichotomy was established at the extended abstract level, thus discriminating between those subjects who adequately went beyond the information presented in the original, and those who did not.

Furthermore, the taxonomy also provided a way of establishing various levels of quality amongst those subjects who did not satisfy the criterion set. This is particularly important in the case of foreign language reading assessment, where it is

only reasonable to expect a wide range of levels of quality of reading outcomes owing to the individual subject's changing - and the group of subjects' differing - level of proficiency.

Another aspect of criterion-referenced evaluation that has been explored in this study, is its relationship with norm-referenced testing. Davies has pointed out that criterion referencing

can only be done after norm referencing. After a test has been shown to discriminate on a population then criterion referencing may be used on the same items in order to provide an appropriate achievement test for a particular sample.

(Davies, 1977a.,p.8)

In spite of the exploratory nature of the study, it seems to me that it provided evidence, in this respect, of the convenience of beginning by establishing as clearly defined a criterion-referenced measure as possible, and then applying standard norm-referenced procedures in the running of the experiment and in the analysis of its results in order to find out whether the criterion set a priori discriminates in practice on the population sampled or not.

In the case of the present study this was done by relying on the SOLO Taxonomy as a basic framework for establishing the required criterion, and on the ANOVA test as the main tool for the statistical analysis of the results of the experiment.

Summing up, it could be said that the research tools were adequate for the purpose of the study regarding the methodological requirements that the use of criterion-referenced and norm-referenced procedures imposed upon it.

6.3.3.3 Hypothesized and Observed Results

In this section the adequacy of the research tools used will be considered in relation to the notion of construct validity.

As Savignon (1983, p.238) points out, "the construct validity of tests of L2 proficiency is...a major issue in current L2 test research." Regarding the nature of construct validity, her view coincides with those that can be found in Cronbach, 1961, p.106 (cited in Davies, 1977b, p.59); Davies, 1977b, pp.62-3; Ingram, 1977, p.22; Kline, 1986, pp.6-7; and Martuza, 1977, pp.150-8. A clear statement of what is involved in the evaluation of the construct validity of a given test is provided by the American Psychological Association:

Construct validity is evaluated by investigating what qualities a test measures, that is, by determining the degree to which certain explanatory concepts or constructs account for performance on the test. To examine construct validity requires a combination of logical and empirical attack. Essentially, studies of construct validity check on the theory underlying the test. The procedure involves three steps. First, the investigator inquires: From this theory, what hypotheses may we make regarding the behavior of persons with high and low scores? Second, he gathers data to test these hypotheses. Third, in light of the evidence, he makes an inference as to whether the theory is adequate to explain the data collected. If the theory fails to account for the data, he should revise the test interpretation, reformulate the theory, or reject the theory altogether. Fresh evidence would be required to demonstrate construct validity for the revised interpretation.

(American Psychological Association, 1966, p.13, cited in Martuza, 1977, p.151).

As far as the present study is concerned, the explanatory concepts or constructs under consideration were "Language of Test," "Level of Proficiency" and "Order of Administration." (Section 3.2.1). A number of hypotheses were formulated regarding the way in

which these variables may affect the quality of reading outcomes. (Section 3.2.3). And, data was collected (Section 3.5) and analysed (Chapter 4) to test these hypotheses. In the light of the evidence it was inferred that the theory was adequate to explain the data collected (Chapter 5). This can be taken as evidence of the construct validity of the experiment. In other words, the fact that it was possible to differentiate various levels of quality amongst the subjects' reading outcomes as a function of the explanatory variables under consideration provides evidence of the construct validity of the study.

It follows, then, that the research instruments used were adequate in terms of the main purpose of the study, i.e., to find out the extent to which the experimental results would match those previously hypothesized. The SOLO Taxonomy was adequate for analysing the quality of the subjects' reading outcomes, and the ANOVA test for establishing the statistical significance of the differences observed.

6.4 Summary of Main Findings

The results of the application of the appropriate ANOVA tests (as reported in Chapter 4, and discussed in Chapter 5), to the data, permitted us to make the following observations regarding the way in which the independent variables ("Language Used in the Test," "Level of Proficiency in the Foreign Language" and "Order of Administration of the Tests") affected the dependent one ("Reading Outcomes") in the present study.

The first effect to be considered will be that of the variable "Order of Administration of the Tests." The reasons for dealing

with it first have been previously discussed; namely, that it was introduced in the study out of practical considerations regarding the type of research design chosen rather than owing to an intrinsic interest in it (cf. Sections 2.3.4, 3.2.1 and 5.3.1), and because - in the event - it was the only one that did not interact significantly with the other two (cf. Sections 4.4.1 and 5.3.1).

Briefly stated, it was found that the subjects' performance improved overall in the second test they took regardless of the language of the test or of their proficiency level. This finding permitted us to reject the null hypothesis of no difference in reading outcome as a function of order of administration of the tests, and in the context of the present study, it was taken as evidence of a learning effect in operation.

Since it was found that the variables "Language Used in the Test" and "Level of Proficiency in the Foreign Language" interacted significantly, it was necessary to analyse the way in which their interaction affected the dependent variable, rather than to examine their separate effects (Section 4.4.1). The results of the analyses performed at each level of the variables under consideration (Section 4.4.2), were discussed in relation to a combined version of the corresponding null hypotheses (Section 5.3.2). This revised null hypothesis was rejected, and consequently, the combined alternative one accepted. In other words, a significant difference in reading outcomes between the mother tongue and the foreign language was observed in favour of the mother tongue, but the difference grew smaller as the subjects' level of proficiency in the foreign language increased. A summary

now follows of what this meant for each of the variables involved.

In the case of the analysis that focused on the component "English" of the variable "Language of the Test," it was found that the higher the subjects' level of proficiency, the higher the quality of their reading outcome. In contrast, no differential effect was found in the case of the component "Spanish" of the same variable (Sections 4.4.2.1, and 5.3.2).

Regarding the analyses that focused on the three components of the variable "Level of Proficiency" (Section 4.4.2.2), the following observations could be made (cf. Section 5.3.2).

As far as the group of "Advanced" subjects is concerned, it was found that the standard of their reading outcomes did not differ in the English and Spanish tests. On the other hand, the "Intermediate" group performed differently in the English and Spanish tests, the difference being in favour of the Spanish test. Finally, the group of "Beginners" was also found to perform better in the Spanish than in the English test, only now the difference was larger than in the case of the "Intermediate" group.

It may be convenient to remember here that a closer look at the way in which the interaction "Language of Test x Level of Proficiency" affected the quality of the subjects' reading outcomes was taken from a SOLO perspective, and that the detailed results of such examination are included in Chapter 5 (Sections 5.2.1 and 5.2.2.2).

Going back to the results summarized above, it might be argued that they do nothing but confirm something that was already

obvious. However, this point is debatable. And also - even if that were the case - it need not be considered as something negative. Let me support these remarks.

Lazarsfeld has pointed out that there is something wrong with the argument of "obviousness" regarding social science research findings. After providing illustrations to support his view, he states:

Since every kind of human reaction is conceivable, it is of great importance to know which reactions actually occur **most** frequently and under what conditions; only then will the more advanced social science develop. (My emphasis).

(Lazarsfeld, 1949, pp.379-80; cited in Jones and Gerard, 1967, pp.34-5)

Let us consider, for the sake of argument, the possibility of having found different results. For example, if the "Order of Administration of Tests" effect had not been significant, it may have been concluded that the learning effect had been kept under perfect control in the experiment. And if on the other hand, it had been found to interact significantly with one of the other two variables, again, an explanation could no doubt have been found. For example, if it had interacted significantly with the "Language of Test" variable, it could have been argued that the "Order of Administration of Tests" had intervened facilitating (or in its case, making it harder) taking a test in a given language (English or Spanish) in a given order (First or Second) so overwhelmingly that all the experimental subjects had been affected that way regardless of their level of proficiency. And, depending on which of the two languages had happened to be more difficult in a given order of administration, a different explanation would have been needed to account for such a finding.

But then, that is not what was actually observed in the present study.

On the other hand, it was also stated above that even if the experimental results do coincide with commonsense observations, it need not be considered as a fault in the experiment. As a matter of fact, as was pointed out in Section 6.3.3.3, for an experiment to have construct validity, the observed and hypothesized results should be clearly related. Otherwise, either the underlying theory has failed to account for the observed results, or the research tools have not been adequate enough to probe the data as required by the purposes of the study, namely, to put to the test the experimental hypotheses. These, it must be emphasized, are nothing else but formal statements of the questions the researcher wishes to find answers to, questions which in turn are based on "commonsense" observations and/or the findings of previous research in the area. So, there is nothing wrong if the experimental results and commonsense statements about the subject matter of a study coincide.

In consequence, it could be said that the usefulness of putting the experimental hypotheses to the test (provided this is properly done) lies in the fact that the results - by rejecting the null form and accepting the alternative one, or otherwise - will contribute to further develop the understanding of the phenomenon under consideration.

Accordingly, it could be argued that the present study served the purpose of providing some empirical evidence about the way in which the variables "Language of Test," "Level of Proficiency" and "Order of Administration" affect the quality of "Reading

Outcomes." The specific relationship observed between these variables can be summarized by the equation in Figure 6.1.

$$(\text{"Language of Test"} \times \text{"Level of Proficiency"}) + \text{"Order of Administration"} = \text{"Reading Outcome," i.e., SOLO Levels}$$

Figure 6.1 Relationship Observed Between the Variables in the Study.

These observations still require further research on other sets of data in order to be cross-validated.

6.5 Implications

Regarding the discussion of possible implications of the study, a practical perspective will be taken first (Section 6.5.1), and then, future lines of research will be suggested (Section 6.5.2).

6.5.1 Practical

An important contribution that criterion-referencing can make to education in general has been briefly expressed by Farr, Carey and Tone (1986, p.148) when they state that "it requires educators to define their goals, objectives, and philosophies." Biggs and Collis (1982) illustrated how their SOLO Taxonomy can be applied to achieve this end in a number of content areas, and Hutchinson has noted in relation to the specific case of language instruction:

There is clearly a need for experts to define more clearly their concept of attainment in language in qualitative terms, and perhaps, a scheme along the lines of Biggs' SOLO Taxonomy would be helpful in assessing responses to more open-ended questions, where candidates are required to generate an original stretch of language which makes explicit their understanding of and reaction to the written text.

(Hutchinson, 1985, p.63).

That this can be done is supported by the results of the present study, which provide evidence of the applicability of the SOLO Taxonomy to foreign language (as well as first language) reading assessment.

But besides a direct application that the results of this study may have on testing, they can be relevant for teachers and materials designers as well.

As mentioned earlier (Section 6.4), it was possible in the present study to discriminate between those subjects who satisfied the criterion set in advance, and those who did not; but also it was also possible to differentiate various levels of quality among the reading outcomes of those who failed.

It was pointed out that having done so was important as far as reading assessment in general was concerned, but that it was particularly so in the case of foreign language reading evaluation, where a wide range of variation can be expected as a result of the varying degree of command of the foreign language.

But the analysis of reading outcomes along the lines provided by a marking scheme of this nature can also be profitable for teachers, students and materials designers as well.

Teachers and students can explore together their reading performance, taking as a point of departure a clearly defined criterion

against which to compare in terms of quality their actual reading outcomes. By examining the adequacy - or otherwise - of their reading outcomes (and/or of the criterion set), they can reach a better understanding of what it is that makes it easy/difficult to reach different levels of complexity in different reading tasks.

As far as materials design is concerned, the main implication of the study is related to the general pedagogic notion of grading. Grading tasks in terms of increasing complexity, and analysing the reading materials to be used to ensure their adequacy for the tasks being required, can be done along the lines proposed by Biggs and Collis and explored in the present study. (As a matter of fact, on completion of the study, I began the preparation of a course aimed at improving the quality of the students' reading in English as a foreign language).

Finally, let me point out that the results of this study support the view "that some threshold [level of competence in the foreign language] does...appear to be necessary before other abilities, like one's first-language reading ability, can be brought to bear upon the task of reading in a foreign language." (Alderson, 1984, p.19). Accordingly, it is unrealistic to expect groups of students with low proficiency levels to attain high quality levels in complex reading tasks. For example, in the case of the specific educational context where this study was carried out, it is clear that reading courses can achieve better results if the students taking them have a level of command of English similar to that of the "Intermediate" subjects of the study. On the other hand, if subjects with a lower level of command

are to be accepted, then the tasks set should be graded accordingly, and the expectations lowered to avoid disappointments.

6.5.2 Future Research

There is a number of possible lines of research that could be taken up to continue exploring the relationship between level of proficiency and quality of reading outcomes along the lines followed in the present study. Some will now be suggested.

It would be convenient to attempt to duplicate the results of this study on a fresh set of data. If it were not possible to do so, it would be interesting to examine the reasons behind the differences that might be observed between the two studies. On the other hand, the closer their findings, the more conclusive the evidence thereby provided about the proficiency-quality of outcome relationship.

It would also be convenient, besides attempting to cross-validate the study in terms of the hypotheses set in it, to examine in depth some patterns that became apparent while analysing its results from a SOLO perspective (Section 5.2). For example, it was observed that limited linguistic resources did not prevent some subjects from reaching the extended abstract level (Section 5.2.1.2). Given the importance of this finding, it would be necessary to submit it to experimental analysis.

Another important line of research that could be suggested is related to the improvement of the marking scheme used in the study. It was pointed out (Section 4.3.2), for example, that although a high degree of interjudge reliability was observed,

there were cases where the two judges who marked the subjects' scripts were in disagreement. It would be profitable to go back to the data and examine in detail the reasons behind their disagreement when it was of more than one level of difference. This examination would be expected to allow a better understanding of what can be judged as determining outcome quality when the reading tasks set require essay-type responses. As a by-product of these considerations, cues might be found as to factors that diminish the reliability of marking schemes of this nature in general. Again, the findings of the proposed analysis should be submitted later to experimental verification.

In a related line of research, the scripts of all the subjects (and not only of those where major differences between the judges were found) could be re-examined in an attempt to identify specific factors that diminish the quality of the subjects' responses besides those regarding their structure (since these were already considered when defining the marking scheme used in the study). For example, it was observed while marking the subjects' responses in this study that some were so general that they could hardly be "wrong," yet, they were not good enough for reaching high quality levels. In this respect, it would be interesting to find out more about what levels of generalization can be considered appropriate for different tasks, and for different levels.

Finally, other lines of future research that can be suggested could expand the scope of the present study by trying out different research methodologies (e.g., by incorporating interviews to the subjects), or by examining other related variables (e.g., different types of reading tasks).

APPENDICES

APPENDIX ATest A(Used in Exploratory Study)

Appendix A Test A(Used in Exploratory Study)

The instructions were exactly the same in Tests A and B, and in both, they were given in Spanish. A translation is given below.

INSTRUCTIONS

Write a summary of the following text. Write it with the purpose of using it to inform fellow students of yours - that may be interested in the topic, but who have not had the chance to read this article - of its content.

Afterwards, if you have them, write your own comments on the material read.

Write both summary and comments in Spanish.

INSTRUCCIONES.

Escriba usted un resumen del siguiente texto. Escribalo con el propósito de exponer su contenido a compañeros de clase que puedan interesarse por el tema pero que no hayan tenido oportunidad de leer este artículo.

Posteriormente, en caso de tenerlos, incluya también sus propios comentarios sobre el material leído.

Escriba usted su resumen y sus comentarios en español.

ECOLOGISTS: POLLUTION SPAWNING NEW ILLNESSES

Have a smoker's cough though you don't smoke? Itchy eyes? Throat and lungs perpetually sore and nose stuffed? Is your sex drive decreased? Do you have slowed reflexes or recurring bad dreams?

You may have what doctors are starting to call "chilango neurosis."

This is one of the new diseases, caused by the unique mix of polluting particles in Mexico City, that are being identified by Dr. Roberto Gomez Navarrete and a group of 30 doctors affiliated with the Mexican Ecological Movement (MEM).

Victims of chilango neurosis experience dramatic mood-swings from elation to depression for no apparent reason. They lose interest in life, get the "couldn't care less" attitude and, in the some cases, try to escape through drugs, alcohol or suicide.

Gomez said chilango neurosis (chilango is slang for Mexico City resident) is caused by air-borne toxic gases that enter the nervous system and affect cerebral functions. The resulting organic damage directly affects behaviour.

Another new illness, according to Gomez, is the permanent cold suffered by many Mexico City residents, which the doctors have dubbed the "toxic cold."

Victims develop a permanently irritated throat and a stuffy nose, symptoms that may worsen and become a cold five or six times a year.

The third common pollution-related affliction is the "metropolitan headache." Dr. Gomez said brain surgeons can treat but not cure this illness. "The only way to get rid of it is to leave the city."

As bad as these diseases are, the ecologists cite even more disquieting health dangers posed by the city's heavy pollution.

Foremost among these is lead poisoning. Merely by breathing, residents are accumulating so much lead in their bodies, and passing it on to the children they bear, that "we are in danger of forming a generation of idiots by the year 2000," according to Alfonso Cipres Villarreal, head of the MEM.

"Over 100,000 children die a year because of pollution," estimates Gomez, president of the Mexico State MEM.

"Lead is one of the principal enemies of those who live in polluted areas," Gomez said. "Research groups...calculated that over 1 million children have dangerous levels of lead in their blood." The researchers come from the National Autonomous University of Mexico, National System for Family Integral Development (DIF) and the National Polytechnic Institute.

He said that symptoms of lead poisoning are memory loss, slowed physical and mental reflexes, lowered sex drive and, in the advanced stages, mental retardation.

Homero Aridjis, creator of the Group of 100, an ecological group composed of 100 prominent Mexican intellectuals and artists, has more to say about lead.

"Lead poisoning causes dizziness, impaired speech and hearing, confusion and, in severe cases, cerebral disfunction," he said, adding that those most susceptible to lead poisoning are children, pregnant women and old people.

Long-time Mexico City residents eventually develop a metabolic process that expels through the mouth, urine and sweat some of the absorbed lead.

But Gomez said that since a foreigner "doesn't have self-defense or adaptive mechanisms, 20 days in Mexico City are worse for him than a year for a permanent resident."

Though Mexico City is believed to have more lead in its air than any other city on earth, no one knows how much there is. Cipres Villarreal estimates 10 tons a day are spewed into the air, while Sedue puts the figure at 18 tons.

Among the chief sources of lead are the Azcapotzalco oil refinery in northern Mexico City and the millions of cars that circulate through the city daily.

Along with lead, vehicles expel 2.5 million tons of carbon monoxide a year, 300,000 tons of hydrocarbons and 30,000 tons of nitrogen oxide, according to Aridjis.

The National Institute of Respiratory Illnesses says 10 percent of the Mexico City population has chronic asthma. There is a very high incidence of conjunctivitis, known as "pink-eye", a contagious infection of the eye membrane. The Arista clinic near the Revolution subway station treats 200 to 250 people daily for eye infections. Last winter, when air pollution reached unprecedented levels in the city, a government official said that seven of every 10 city residents were suffering respiratory afflictions.

Food is another source of danger cited by the ecologists. Crates of tomatoes unloaded at the supermarket are stamped "For national consumption. Export prohibited." The pesticides with which they have been sprayed are illegal in other countries.

There are 260 pesticides used on fruits and vegetables Mexico City residents eat. The Secretariat of Agriculture says only five of these are authorized.

A recent United Nations World Health Organization report said food in the Mexico City area is very high in pesticides, and estimated that each year, 500,000 people suffer from food poisoning.

The excessive noise levels in the urban area can cause hearing loss, neuroses and stunted growth in children, according to the Group of 100. It affects kidneys and blood vessels.

According to the group, 68 percent of the gas motors and 87 percent of diesel motors produce over 85 decibels of noise, the limit considered acceptable.

Though little research has been done on them, pungent odors in the city also threaten residents' health. The odors, such as those emitted on by a detergent factory in Colonia Irrigacion, cause appetite loss, nausea, respiratory problems, cardiovascular complications and allergies, according to the Group of 100.

Cipres Villarreal said the government hides information from the public regarding the severity of the city's pollution.

"The knowledge could cause social movements which would alter the order of Mexican society," Cipres Villarreal said.

President Miguel de la Madrid created the Secretariat of Urban Development and Ecology (Sedue) at the beginning of his term in 1982. But the government has compiled almost no statistics on the effects pollution has on the health of capital residents.

Patricia Gilhuys, a public relations officer at Sedue, said they had no concrete statistics on health hazards posed by pollution, and told a reporter to phone the Health Secretariat.

Alfredo Valle, a Health Secretariat press officer, said there were no official studies, but promised to "pull something out of the archives."

By Heather Evans and Peter Kulten-
brouwer.

The News Staff Reporters.

The Mexico City News.

Tuesday, May 20, 1986.

APPENDIX BTest B(Used in Exploratory Study)

Appendix BTest B(Used in Exploratory Study)

Please see the note about the instructions in both Tests A and B, that is included at the beginning of Appendix A.

ECOLOGOS: LA CONTAMINACION FUENTE DE ENFERMEDADES NUEVAS.

¿Tiene usted tos de fumador aunque no fume? ¿Ardor en los ojos? ¿La garganta y los pulmones perpetuamente irritados y la nariz tapada? ¿Ha decaído su impulso sexual? ¿Tiene reflejos retardados o pesadillas recurrentes?

Tal vez tenga usted lo que los médicos han dado en llamar "la neurosis chilanga".

Esta es una de las nuevas enfermedades que, causadas por la mezcla singular de partículas contaminantes en la Ciudad de México, han sido identificadas por el Dr. Roberto Gómez Navarrete y un grupo de 30 médicos afiliados al Movimiento Ecológico Mexicano (MEM).

Las víctimas de la "neurosis chilanga" sufren cambios bruscos de estado de ánimo, que van de la euforia a la depresión, sin una razón aparente. Pierden interés por la vida, adoptan una actitud de "a mí qué me importa", y en algunos casos, tratan de escapar por medio de las drogas, el alcohol o el suicidio.

El Dr. Gómez indicó que la "neurosis chilanga" ("chilango" se usa en el habla popular para referirse a los habitantes de la Cd. de México) es causada por gases tóxicos presentes en el aire que entran al sistema nervioso y afectan las funciones cerebrales. El daño orgánico resultante afecta directamente la conducta.

Otra enfermedad nueva, de acuerdo al Dr. Gómez, es el resfriado permanente que sufren muchos habitantes de la Cd. de México, y que los médicos han dado en llamar "el resfriado tóxico".

Las víctimas permanentemente tienen la garganta irritada y la nariz tapada, síntomas que pueden empeorar y convertirse en resfriado cinco o seis veces al año.

La tercera aflicción común relacionada a la contaminación es el "dolor de cabeza metropolitano". Señaló el Dr. Gómez que los especialistas en padecimientos cerebrales pueden tratar, pero no curar esta enfermedad. "La única forma de evitarla es irse de la ciudad".

Aunque estas enfermedades son en sí graves, los ecólogos citan otros peligros para la salud aún más intranquilizantes que son ocasionados por la fuerte contaminación.

El principal de éstos es el envenenamiento por plomo. Por el sólo hecho de respirar, los habitantes de la ciudad están acumulando tanto plomo en el cuerpo, y transmitiéndolo a sus hijos, que "estamos en peligro de formar una generación de idiotas para el año 2000", según opinión de Alfonso Ciprés Villarreal, director del MEM.

"Más de 100,000 niños mueren al año debido a la contaminación", calcula el Dr. Gómez, presidente del MEM en el Estado de México.

"El plomo es uno de los enemigos principales para quienes viven en áreas contaminadas", dijo el Dr. Gómez. "Algunos grupos de investigadores...han calculado que más de un millón de niños alcanzan niveles peligrosos de plomo en la sangre". Estos investigadores provienen de la Universidad Nacional Autónoma de México, el Sistema Nacional para el Desarrollo Integral de la Familia (DIF) y el Instituto Politécnico Nacional.

Indicó el Dr. Gómez que los síntomas del envenenamiento por plomo son pérdida de la memoria, deterioro en los reflejos físicos y mentales, disminución del impulso sexual y, en las etapas avanzadas, retraso mental.

Homero Aridjis, fundador del Grupo de los 100, un grupo ecológico formado por 100 destacados intelectuales y artistas mexicanos, tiene algo que añadir acerca del plomo.

"El envenenamiento por plomo causa mareos, deterioro en el habla y el oído, confusión y, en casos severos, disfunción cerebral", dijo, añadiendo que los más susceptibles al envenenamiento por plomo son los niños, las mujeres embarazadas y los ancianos.

Los habitantes de la Cd. de México que tienen mucho tiempo de radicar en ella, llegan a desarrollar un proceso metabólico que expulsa por el aliento, la orina y el sudor parte del plomo absorbido.

Sin embargo, señaló el Dr. Gómez, puesto que un extranjero "no tiene mecanismos de defensa o de adaptación, 20 días en la Cd. de México son peores para él que un año para un residente permanente".

Aunque se piensa que la Cd. de México tiene más plomo en el aire que cualquier otra ciudad en el mundo, nadie sabe cuánto hay. Ciprés Villarreal calcula que diariamente se arrojan 10 toneladas al aire, mientras que SEDUE da la cifra de 18 toneladas.

Entre las fuentes principales de plomo están la refinería de petróleo de Azcapotzalco al norte de la Cd. de México y los millones de automóviles que circulan diariamente por la ciudad.

Junto con el plomo, los vehículos arrojan 2.5 millones de toneladas de monóxido de carbono al año, 300,000 toneladas de hidrocarburos y 30,000 toneladas de óxido de nitrógeno, según Aridjis.

El Instituto Nacional de Enfermedades Respiratorias dice que el 10 por ciento de la población de la Cd. de México padece asma crónica. Hay una alta incidencia de conjuntivitis ("ojos rojos"), infección contagiosa de la membrana ocular. La clínica Arista, cercana a la estación Revolución del metro, atiende diariamente de 200 a 250 personas que sufren de infecciones en los ojos. El invierno pasado, cuando la contaminación ambiental alcanzó niveles sin precedente en la ciudad, un funcionario del gobierno dijo que siete de cada 10 habitantes de la ciudad sufrían de algún padecimiento de las vías respiratorias.

La comida es otra fuente de peligros citada por los ecólogos. Hay cajas de jitomates que se descargan en los supermercados marcadas: "Para consumo nacional. Prohibida la exportación". Los pesticidas con que se han rociado son ilegales en otros países.

Se utilizan 260 pesticidas para fumigar las frutas y verduras que comen los habitantes de la Cd. de México. La Secretaría de Agricultura señala que solamente cinco están autorizados.

En un informe reciente de la Organización Mundial de la Salud de las Naciones Unidas se indicó que la comida en el área de la Cd. de México contiene niveles muy altos de pesticidas, y se calculó que cada año 500,000 personas sufren de intoxicaciones por su ingestión.

De acuerdo al Grupo de los 100, los niveles excesivos de ruido en el area metropolitana pueden causar pérdida del oído, neurosis y falta de crecimiento en los niños. Los riñones y los vasos sanguíneos también se ven afectados.

De acuerdo al grupo, 68 por ciento de los motores de gasolina y 87 por ciento de los motores diesel producen más de 85 decibeles de ruido, el límite que se considera aceptable.

Aunque existe poca investigación sobre los malos olores en la ciudad, también éstos constituyen una amenaza para la salud. Según el Grupo de los 100, dichos malos olores, como los emitidos por una fábrica de detergentes en la Colonia Irrigación, causan pérdida del apetito, náuseas, problemas respiratorios, complicaciones cardiovasculares y alergias.

Ciprés Villarreal opinó que el gobierno oculta información al público sobre la gravedad de la contaminación de la ciudad.

"Esta información podría provocar movimientos sociales que alterarían el orden en la sociedad mexicana", añadió Ciprés Villarreal.

El presidente Miguel de la Madrid creó la Secretaría de Desarrollo Urbano y Ecología (SEDUE) al principio de su gestión en 1982. Sin embargo, el gobierno no ha recopilado casi nada de datos estadísticos sobre los efectos de la contaminación en la salud de los habitantes de la capital.

Patricia Gilhuys, una encargada de relaciones públicas en SEDUE, dijo que se carece de estadísticas concretas sobre los riesgos que la contaminación conlleva para la salud, e indicó al periodista que llamara por teléfono a la Secretaría de Salud.

Alfredo Valle, un oficial de prensa de la Secretaría de Salud señaló que no habían estudios oficiales, pero prometió "sacar algo de los archivos".

Por Heather Evans y Peter Kulten-
brouwer.

Periodistas de The News.

The Mexico City News.

Martes 20 de mayo de 1986.

APPENDIX C Transcripts of Responses Obtained in an
Exploratory Application of Tests A and B

Appendix C Transcripts of Responses Obtained in an
Exploratory Application of Tests A and B

The transcripts of the responses given in Tests A and B by three subjects ("A," "B" and "C") are included in this appendix.

SUBJECT: A.

TEST: A. (Text in English).

POLUCION

LA CREACION DE UNA NUEVA ENFERMEDAD.

¿Toses sin ser fumador?; ¿Sientes molestias en los ojos?; ¿Padeces de la garganta o de los pulmones?; ¿Está tu actividad sexual disminuyendo?; ¿Tienes reflejos lentos o duermes mal?

Estas son las preguntas que le han hecho un grupo de doctores mexicanos a la población. Estos doctores están dirigidos por Don Roberto Gómez Navarrete. Ellos exponen el problema de la polución al mismo nivel que el problema de la droga, alcoholismo o el alto nivel de suicidios actualmente.

Para ellos la contaminación atmosférica afecta a la temperatura ambiental y posiblemente a causa de ello la temperatura es mas baja que años atras. Otro problema es la alta concentración de ruido con la posible consecuencia de pérdida de oído. A esta concentración de ruidos los doctores le llaman "dolor de cabeza ciudadano".

Según un estudio hecho todo esto puede formar una generación tullida para el año 2000. Según una estadística cerca de 1.000.000 de niños están en grave peligro y 100.000 mueren cada año a causa de todos los problemas expuestos anteriormente.

Los síntomas más acusados son la pérdida de memoria; lentos reflejos físicos y mentales y atraso mental.

Según el Señor Roberto Gómez Navarrete para un extranjero no residente en México Capital, el organismo activa las defensas orgánicas o sufre un proceso de adaptación progresivo.

Según el Instituto Nacional de Enfermedades Respiratorias un 10% de la población de México Capital sufre de asma crónica y otras enfermedades relacionadas con la polución.

Otro problema producto de la polución y que afecta directamente a la salud pública es referente a los alimentos, los cuales han sido tratados con abonos no permitidos por la ley. Según un estudio, los vegetales, hortalizas y frutas son tratados con 260 pesticidas de los cuales sólo 5 están autorizados. Y a causa de esto 500.000 personas sufren o ingieren cierta cantidad de veneno cada año.

Otro problema expuesto anteriormente es el alto nivel de ruido producido por el quehacer cotidiano. Esto es otro modo de contaminación ambiental tan grave como los ya mencionados anteriormente.

¿Puede provocar movimientos sociales el conocimiento de las causas, efectos y consecuencias de la contaminación? Esta es una pregunta con difícil respuesta. Para evitar esto el gobierno ha creado unas estadísticas y sus posibles soluciones.

SUBJECT: A.

TEST: B. (Text in Spanish).

LA CONTAMINACION FUENTE DE ENFERMEDADES NUEVAS.

¿Tiene usted tos de fumador aunque no fume? ¿Ardor en los ojos? ¿La garganta y los pulmones perpetuamente irritados y la nariz tapada? ¿Ha decaído su impulso sexual? ¿Tiene reflejos retardados o pesadillas recurrentes?

Tal vez tenga usted lo que los médicos han dado en llamar "la neurosis chilanga".

Estas son las preguntas que se han hecho un grupo de científicos afiliados al Movimiento Ecológico Mexicano (MEM), a fin de identificar los síntomas de dicha enfermedad.

Para este grupo de científicos los síntomas más aparentes de la llamada "neurosis chilanga" se manifiestan principalmente en el sistema nervioso y afectan las funciones cerebrales y en general la conducta del individuo. Todo esto está producido por la enorme cantidad de gases tóxicos que aspiramos cada día. Estos gases tóxicos al afectar al sistema nervioso y vías respiratorias producen el llamado "resfriado tóxico" y el "dolor de cabeza metropolitano".

El principal agente contaminante es el plomo, ya que es un agente contaminante que se va acumulando y se puede hasta transmitir a los hijos y según el MEM "estamos en peligro de formar una generación de idiotas para el año 2000". Más de 100.000 niños mueren al año por causas de agentes contaminantes y cerca de 1.000.000 de niños alcanzan

niveles peligrosos de plomo en la sangre. Según unos investigadores los síntomas principales por envenenamiento por plomo son: pérdida de memoria, deterioro en el habla y el oído y confusión mental. La población más afectada por estos agentes contaminantes son los niños, las mujeres embarazadas y los ancianos.

Los habitantes de la Cd. de México llegan a desarrollar un proceso metabólico que expulsa por el aliento, la orina y el sudor parte del plomo absorbido. Pero para un extranjero que no tiene mecanismos de defensa o de adaptación, 20 días en la Cd. de México son peores para el que un año para un residente permanente.

Junto al plomo hay otros agentes contaminantes, como son el monóxido de carbono, hidrocarburos y óxido de nitrógeno.

Según fuentes oficiales un 10% de la población de la Cd. de México padece asma crónica. En el invierno pasado, cuando la contaminación alcanzó su punto máximo, un funcionario del gobierno dijo que siete de cada 10 habitantes de la ciudad sufrían de algún padecimiento de las vías respiratorias.

Los alimentos son otro problema. Se utilizan 260 pesticidas para fumigar las frutas y verduras que comen los habitantes de la Cd. de México, de los cuales sólo 5 pesticidas están autorizados por la ley. Y a causa de esto cada año 500.000 personas sufren de intoxicaciones por su ingestión.

Otro modo de contaminación es el ruido producido en el ajetreo diario.

Según el MEM el gobierno oculta esta información al público, ya que esto podría provocar movimientos sociales que alterarían el orden en la sociedad mexicana.

El gobierno prometió buscar soluciones pero desde 1982 hasta nuestros días no se ha buscado ninguna solución efectiva al problema. Un oficial de prensa de la Secretaría de Salud señaló que no habían estudios oficiales, pero prometió "sacar algo de los archivos".

SUBJECT: B.

TEST: A. (Text in English).

Ecologistas:

La Polución desarrolla nuevas enfermedades.

Por: Heather Evans y Peter Kultenbrouwer.

(Los reporteros del equipo de noticias).

El artículo desarrolla una serie de comentarios a entrevistas hechas a varias personalidades que tienen que ver con la polución en la Ciudad de México, ya sean oficiales, como miembros de la SEDUE (Secretaría de Desarrollo Urbano y Ecología), la Universidad Nacional Autónoma de México, el Sistema Nacional para el Desarrollo Integral de la Familia, y hasta el Presidente de México, Miguel de la Madrid. Junto con la opinión de varios médicos y miembros del "Grupo de los 100", un grupo ecologista.

Estos últimos hacen un llamado a la opinión pública y critican a las autoridades acerca de la gravedad que alcanzan los niveles de polución en la Ciudad de México. Y mencionan los diferentes tipos y sus consecuencias:

- Los gases tóxicos, que al entrar al sistema nervioso afectan las funciones cerebrales, alterando la conducta, causando estados depresivos y apáticos que conducen al uso de drogas, alcohol o al suicidio.

- Esto causa también enfermedades respiratorias, que van desde el resfriado común hasta el asma, pasando por dolores de cabeza llamado "el dolor de cabeza metropolitano".

- Uno de los gases mas tóxicos de éstos, es el plomo. Producido por la refinería de Azcapotzalco y los millones de autos que circulan en la ciudad. Este tipo de envenenamiento causa pérdida de la memoria, lentitud motriz, reflejos lentos, falta de apetito sexual, y en estados avanzados, retraso mental.

- También la comida constituye un peligro. Al ser tratada con pesticidas que son ilegales en el país, causando con esto envenenamiento por comida.

- Por último, está la contaminación por ruido, pues los niveles que alcanza éste en la ciudad rebasa los normales aceptables por el ser humano.

Además, éstos comentan que existiendo estos problemas, no hay suficiente investigación al respecto, ni una respuesta oficial satisfactoria al asunto.

SUBJECT: B.

TEST: B. (Text in Spanish).

Ecólogos: La Contaminación, Fuente de Enfermedades Nuevas.

Por: Heather Evans y Peter Kultenbrouwer.

Periodistas de "The News. The Mexico City News".

Martes 20 de mayo de 1986.

RESUMEN:

Los reporteros hacen una serie de comentarios sobre el grado de contaminación en la Ciudad de México, y la serie de enfermedades que esto conlleva.

Una de las primeras enfermedades que se mencionan es la "neurosis chilanga", en la que, a causa de la mezcla de gases que respiran los habitantes de la ciudad, se ve afectada la conducta.

También hay daños orgánicos causados por estos gases tóxicos, que van desde dolores de cabeza a resfriados continuos.

Estos en sí no son tan graves como los daños causados por la inhalación de plomo, producido por los millones de autos que hay en el area metropolitana, y la refinería de la ciudad. Estas inhalaciones provocan pérdida de la memoria, lentitud en los reflejos físicos y mentales, disminución del impulso sexual y hasta retraso mental. Afecta principalmente a los niños, viejos y mujeres embarazadas.

Hay otras enfermedades causadas por los gases, como son el asma e infecciones de los ojos.

Otra fuente de peligro es el envenenamiento por comida, ya que las verduras que llegan a abastecer la Cd. de México son regadas con pesticidas que no han sido aprobados por la Secretaría de Agricultura.

Por último, queda la contaminación por ruido, que también es grave, ya que la mayoría de los vehículos producen ruido por sobre el nivel aceptable para el ser humano.

Toda esta información es brindada por organizaciones ecologistas de la Cd. de México, las cuales hacen un llamado al gobierno para que se haga más investigación y se tomen medidas contra esto. Sin embargo, parece que el gobierno aún no se ha ocupado seriamente de este problema.

COMENTARIO:

Me parece bueno el artículo. Las fuentes que cita, confiables. "Creo que artículos como éste ampliarían la conciencia de los capitalinos sobre la gravedad real de la situación en la que viven. Una vez más (aquí y en otras partes del mundo sucede) se ve como los funcionarios y políticos tienen intereses divorciados del bienestar de la población.

El artículo aunque para mí no presenta mucha novedad, sí remueve la preocupación e indignación, lo cual esperamos que conduzca a los lectores a una mayor acción al respecto del problema tratado. Por lo tanto es un buen reportaje, aunque corto.

SUBJECT: C.

TEST: A. (Text in English).

Ecologistas: Gran número de nuevas enfermedades producto de la polución.

Investigadores de varias instituciones: Movimiento Ecologista Mexicano, Universidad Nacional Autónoma de México, Departamento de Integración Familiar, Instituto Politécnico Nacional, El Grupo de los 100 y la Secretaría de Ecología y Desarrollo Urbano, denuncian a la polución de la Cd. de México como fuente de "nuevas enfermedades".

Entre estas enfermedades mencionan: neurosis del chilango, catarro tóxico, migraña metropolitana, y envenenamiento por plomo, asma crónica, conjuntivitis, envenenamiento por comida, etc. Los niños, las mujeres embarazadas y los ancianos son las personas más susceptibles de adquirir estas enfermedades.

Las causas principales de estas afecciones son partículas contaminantes en el aire, pesticidas en frutas y verduras, y el ruido. La refinería de petróleo, las fábricas y los vehículos automotores son los principales causantes de estos contaminantes.

Aunque el gobierno tiene información al respecto, carece de análisis estadísticos, y aún más, esconde esta información al público. "El conocimiento de este problema pudiera causar movimientos sociales, los cuales podrían alterar el orden de la sociedad mexicana".

SUBJECT: C.

TEST: B. (Text in Spanish).

RESUMEN:

Ecólogos: la contaminación fuente de enfermedades nuevas.

Este artículo menciona la denuncia de dos grupos ecologistas, Movimiento Ecologista Mexicano (MEM), y un grupo de 100 intelectuales y artistas mexicanos (Grupo de los 100), en relación a los niveles y efectos de la contaminación en la Ciudad de México. Además, critican la actitud del gobierno con respecto a estos problemas.

Funcionarios de estos grupos mencionan una serie de "nuevas enfermedades", provocadas por los altos niveles de contaminación en la Ciudad de México. Entre estas enfermedades destacan: la neurosis chilanga (chilango es el habitante de la Ciudad de México), el resfriado tóxico, dolor de cabeza metropolitano, envenenamiento por plomo, conjuntivitis, envenenamientos por comida, etc.

El principal contaminante en la atmósfera es el plomo, aunque también mencionan el monóxido de carbono, hidrocarburos, óxido de nitrógeno, pesticidas, el ruido y los malos olores. Las principales fuentes de contaminación son: la refinería de petróleo, los vehículos automotores, diversas fábricas y alimentos contaminados con pesticidas (prohibidos en otros países).

Aunque existen instituciones que investigan algunas de las causas y efectos de la contaminación, Universidad Nacional Autónoma de México (UNAM), Instituto Politécnico

Nacional (IPN), y el Departamento de Integración Familiar (DIF), el gobierno no tiene recopilaciones estadísticas sobre los efectos de la contaminación en la salud de los habitantes. Más aún, el director del MEM opina que las instituciones gubernamentales, Secretaría de Ecología y Desarrollo Urbano (SEDUE), y Secretaría de Salud, etc., ocultan información al público, pues "esta información podría provocar movimientos sociales que alterarían el orden social".

COMENTARIO:

El artículo es un poco sensacionalista y maneja datos y hechos de los cuales "no existen análisis estadísticos de los efectos de la contaminación en la salud".

Sin embargo, es importante denunciar el peligro y magnitud del problema, así como crear grupos de análisis y discusión. El artículo no menciona las actividades difusivas de estos grupos, ni tampoco los posibles mecanismos de aliviar este problema.

APPENDIX D Test 1 (English)

Appendix D Test 1 (English)

The "Instructions," "Summary" and "Comments" sheet are exactly the same for Tests 1 and 2, both in English and Spanish. A translation of the instructions into English is given below.

INSTRUCTIONS

1. Write a SUMMARY of the following text.

Write it with the purpose of informing fellow students of yours - who might be interested in its topic, but who have not had the chance to read it - about its content.

Write your summary in the answer sheet provided for that purpose. (The one marked "SUMMARY").

Write your summary in Spanish.

2. Afterwards, include your own COMMENTS on the material read.

Make sure you write them in the corresponding answer sheet. (The one marked "COMMENTS").

Write your comments in Spanish too.

N.B. You have a time limit of one hour to complete both tasks (summary and comments).

THANK YOU VERY MUCH FOR YOUR COOPERATION.

INSTRUCCIONES

1. Escriba usted un RESUMEN del siguiente texto.

Escribalo con el propósito de exponer su contenido a compañeros de clase que puedan interesarse por el tema, pero que no hayan tenido la oportunidad de leer este texto.

Escriba su resumen en la hoja provista para tal efecto. (La que dice "RESUMEN").

Escriba su resumen en español.

2. Posteriormente, incluya también sus propios COMENTARIOS sobre el material leído.

Asegúrese de escribirlos en la hoja correspondiente. (La que dice "COMENTARIOS").

También escriba sus comentarios en español.

NOTA. Dispone usted de una hora como máximo para terminar ambas tareas (resumen y comentarios).

MUCHAS GRACIAS POR SU COLABORACION.

...Man's impacts on the environment result primarily from growth—growth in his numbers and growth in his appetite for material goods. Both must be brought under control.

Population growth in the industrialized societies has already levelled off; growth-rates in the less-developed parts of the world continue to be high, but are also showing signs of subsiding. Nevertheless it seems possible that, by the early years of the next century, planet Earth will have to provide a home for double the present population—about four fifths of them in the developing world—and this will create enormously increased pressures on natural resources and on the biosphere. Add to this the pressure that will result from the drive of the poor—who make up the major portion of the world's population—to improve their conditions of life, and the tremendous pressures already being placed on the system by the insatiable appetites of the rich minority, and you will see that something has clearly got to give way. In my view, that 'something' must be the appetites of the rich.

Population growth in the developing world must be stabilized, and this will happen either through voluntary methods of population control or the traumatic and tragic consequences of famine and disease or conflict.

Even the most hopeful assumptions about limitation of population growth in the developing countries lead to the conclusion that their claims on world resources will escalate sharply within the next three decades. For quite apart from their increasing numbers of mouths to be fed, they cannot be denied their right to achieve better conditions of life, and surely natural justice dictates that their claims be given top priority—firstly, to assure the meeting of their basic needs for a life that is compatible with human dignity and well-being, and secondly, for equal opportunity to share more fully in the benefits which our technological civilization now makes possible. This can only happen if the industrialized societies reduce the pressures which they are exerting on the biosphere and on natural resources.

There is little sign that the developed countries will reduce their demands voluntarily, for their existing commitment to continuous growth in gross national product is built right into the economic system by which modern industrialized societies function. It is based on the assumption that more is better—that the well-being of the societies can only be assured by continuous growth in the material sense. The expectations of consumers, the creation of employment, the incentives which motivate investors and managers, are all geared to this system in which material growth is the prime mover. And while there has been some serious questioning of this concept of growth during recent times, it remains the dominant theme on which modern societies operate.

To the people of the wealthy countries of the industrialized world, the suggestion that they should reduce their demands on global resources is immediately equated with calls for 'no-growth'—a state which none of them could conceive of as tolerable. Indeed, no-growth in the sense of the cessation of economic growth is not perceived as a viable alternative by any nation: it is equated with stagnation and resulting fears of unemployment, social unrest and reduced standards of living.

It is this attitude towards growth—of human population and its production and therefore demands on the life-support system—that is at the heart of our present dilemma. This is the disease which has spread through the body of modern technological societies. It is this growth disease which has within it the potential for self-destruction of our society; and it is to the curing of this disease that we must direct ourselves if we are to deal fundamentally and effectively with the need to create a sustainable balance between Man and the natural environment.

NOMBRE _____

NIVEL _____ FECHA _____

GRUPO _____ HORARIO _____

RESUMEN

NOMBRE _____

NIVEL _____ FECHA _____

GRUPO _____ HORARIO _____

COMENTARIOS

APPENDIX E Test 2 (English)

Appendix E Test 2 (English)

Please see the note included at the beginning of Appendix D regarding the instructions and answer sheets used in all of the reading comprehension tests.

It is understandable that the early emphasis of environmentalists was on cleaning up some of the more obvious cases of environmental damage, on stopping development which threatened further damage, and on subjecting existing sources of pollution to controls or 'add on' technologies designed to minimize their environmental impacts. Now, some environmentalists have called for a halt to growth in our highly industrialized societies. At the other extreme, there has been a tendency to regard the costs of environmental protection or improvement as simply an extra or added cost—something that cuts into economic return, often to the point where we cannot afford it. We are told that protection of the environment is nice if you can afford it, but that when there is a conflict between economic and environmental factors, economic reality must prevail.

Yet what kind of economic reality is it that leaves out of the cost-benefit calculations (on which economic decisions are based), the entire cost to society as a whole of maintaining the natural environmental capital of air, water, soil, and plant and animal life, on which human life and well-being depend? By what economic logic can the amounts required to preserve environmental values in a given economic situation be considered a 'cost', while the impairment of environmental capital that results from failure to make these expenditures is not seen as a cost? Surely, the reality lies in recognizing that in any activity which damages the environment, environmental costs cannot be avoided—that it is simply a question of how and when and by whom they are to be borne. Environmental realities cannot be avoided by ignoring them in the name of economic needs.

Economic growth has made possible some unprecedented benefits to society. But it has also enabled us to avoid facing up to certain basic truths which we must now confront. We now know that economic growth which depends on the running down and impairment of our natural environmental capital is neither sound nor sustainable. The bills must come in eventually. The growing damage, to our natural environmental capital, that is resulting from the same processes which produce the wealth of our societies, should make it clear to us that we cannot continue to grow as we

have been doing in the past. A society which bases its growth on the degradation and using up of its natural environmental capital, will be no more viable than a business enterprise which does not provide an adequate depreciation and amortization account to maintain its capital and productive capacity. Yet, this is exactly what is happening today in varying degrees in almost all countries in both the industrialized and the developing world!

Individual examples of this wantonness are legion in the fouling of rivers and lakes, contamination of air, desecration of coastal areas, destruction of bird and other animal life, and deterioration of urban areas in the United States and other industrialized countries, while in the developing world we have massive destruction of forests, large-scale loss of productive soil through erosion, march of the deserts and other consequences of ecologically unsound land-use practices, and the appalling conditions of life in 'exploding' urban areas through lack of adequate water supplies, waste disposal and sanitation facilities. There is also the growing threat to rich and poor alike from such 'outer limit' risks as possible climatic change, damage to the ozone shield, contamination of human food-chains, pollution of the oceans and overexploitation of their living resources.

Surely it must be clear that present growth-patterns and practices are self-destructive and cannot be sustained! Is no-growth, then, the only answer? Let me say with all the force I can muster that no-growth is NOT the answer. The real alternative to no-growth is new-growth—a new approach to growth, in both the more industrialized and the less-developed societies.

The new-growth approach must be based upon removal of the artificial and self-defeating conflict between ecology and economics, which is now built into our system of economic decision-making....

APPENDIX F Test 1 (Spanish)

Appendix F Test 1 (Spanish)

Please see the note at the beginning of Appendix D regarding the instructions and answer sheets used in all of the reading comprehension tests.

...El impacto del hombre en el medio ambiente resulta básicamente del crecimiento; crecimiento de su número, y crecimiento de su apetito de bienes materiales. Ambos deben someterse a control.

El crecimiento de la población en las sociedades industrializadas ya se ha nivelado; los índices de crecimiento en las regiones menos desarrolladas del mundo continúan siendo altos, aunque también muestran signos de estar disminuyendo. Sin embargo, parece ser que para los primeros años del próximo siglo el planeta Tierra tendrá que albergar al doble de la población actual -alrededor de cuatro quintas partes en los países en vías de desarrollo- y esto acarreará presiones crecientes sobre los recursos naturales y la biósfera. Añádase a esto la presión que resultará de las demandas de los pobres -quienes constituyen la mayor parte de la población mundial- para mejorar sus condiciones de vida, así como las tremendas presiones que ya se le imponen al sistema por el apetito insaciable de la minoría rica, y se verá que desde luego algo tiene que ceder. En mi opinión, ese "algo" deben ser las ambiciones de los ricos.

El crecimiento de la población en los países en vías de desarrollo debe estabilizarse, y esto ha de suceder sea mediante métodos de control de población voluntarios, o por las consecuencias trágicas y traumáticas del hambre y la enfermedad o el conflicto.

Aun las suposiciones más optimistas en cuanto a la restricción del crecimiento poblacional en los países en vías de desarrollo nos conducen a concluir que sus demandas sobre los recursos mundiales habrán de escalar bruscamente en las próximas tres décadas. Aparte del creciente número de bocas que alimentar, no se puede negar su derecho a alcanzar mejores condiciones de vida, y ciertamente la justicia natural dicta que se dé alta prioridad a sus reclamos -en primer lugar, para garantizar la satisfacción de las necesidades básicas de una vida compatible con la dignidad y el bienestar humanos, y en segundo lugar, de igual oportunidad para gozar más plenamente de los beneficios que nuestra civilización tecnológica hace posibles ahora. Esto puede suceder solamente si las sociedades industrializadas reducen las presiones que están ejerciendo sobre la biósfera y los recursos naturales.

Hay pocos signos de que los países desarrollados vayan a reducir voluntariamente sus demandas, ya que su cometido actual por un crecimiento continuo del producto nacional bruto forma parte del mismo sistema económico con que funcionan las sociedades industrializadas modernas. Se basa en la suposición de que entre más, mejor; de que el bienestar de las sociedades puede asegurarse solamente por el crecimiento continuo en el sentido material. Las expectativas de los consumidores, la creación de empleos, los incentivos que motivan a inversionistas y ejecutivos, todo forma parte de este sistema en que el crecimiento material es la fuerza motriz. Y, aunque recientemente se han expresado algunas dudas serias sobre esta concepción del crecimiento, continúa siendo el tema dominante sobre el que operan las sociedades modernas.

Para la gente de los países ricos del mundo industrializado, la sugerencia de que deberían reducir sus demandas sobre los recursos del mundo es inmediatamente equiparable con una petición de alto al crecimiento, situación que ninguno de ellos podría concebir como tolerable. De hecho, la ausencia de crecimiento en el sentido de un cese del crecimiento económico no se percibe en ninguna nación como alternativa viable: se le equipara con el estancamiento y los resultantes temores al desempleo, la inestabilidad social, y niveles de vida mermados.

Es esta actitud ante el crecimiento (de la población humana y su producción, y consecuentes exigencias al sistema en el que sostiene la vida), la que constituye el meollo de nuestro dilema actual. Este es el mal que se ha extendido en el organismo de las sociedades tecnológicas modernas. Es esta enfermedad del crecimiento la que lleva en sí misma el potencial de la autodestrucción de nuestra sociedad; y es a la cura de este mal que nos debemos abocar si hemos de hacer frente a fondo y eficazmente a la necesidad de crear un equilibrio perdurable entre el hombre y el medio ambiente natural.

APPENDIX G Test 2 (Spanish)

Appendix G Test 2 (Spanish)

Please see the note included at the beginning of Appendix D regarding the instructions and answer sheets used in all of the reading comprehension tests.

Es comprensible que el énfasis inicial de los ecologistas haya estado en la limpieza de algunos de los casos más obvios de daño al medio ambiente, en detener el desarrollo que amenazara con causar más daño, y en someter las fuentes de contaminación existentes a controles o a tecnologías complementarias diseñadas para minimizar sus impactos en el medio ambiente. Actualmente algunos ecologistas han exigido un alto al crecimiento de nuestras sociedades. En el extremo opuesto, ha habido la tendencia a considerar los costos de la protección o el mejoramiento ambientales como simplemente un costo extra o adicional -algo que va en detrimento de la ganancia económica, a menudo a un punto que no podemos permitirnos. Se ha dicho que la protección del ambiente está bien si se puede costear, pero que cuando hay conflicto entre los factores económicos y los ambientales, la realidad económica debe prevalecer.

Sin embargo, ¿qué clase de realidad económica es aquella que excluye de los cálculos de costo-beneficio (en los cuales se basan las decisiones económicas), el costo total que significa para la sociedad entera el mantenimiento del capital ambiental natural de aire, agua, suelo, y vida animal y vegetal, del que dependen la vida y el bienestar humanos? ¿En qué lógica económica cabe considerar como "un costo" las sumas que se requieren para preservar los valores ambientales en una situación económica dada, a la vez que el deterioro del capital ambiental que resulta por no realizar estos gastos no se ve como un costo? Seguramente la realidad estriba en reconocer que en cualquier actividad que dañe al medio ambiente los costos ambientales no pueden evitarse, y que es sencillamente una cuestión de cómo y cuándo y quién ha de sobrellevarlos. Las realidades ambientales no pueden evitarse ignorándolas en nombre de necesidades económicas.

El crecimiento económico ha hecho posibles para la sociedad beneficios sin precedente. Pero también nos ha impedido encarar ciertas verdades básicas que ahora tenemos que confrontar. Hoy sabemos que el crecimiento económico que depende del agotamiento y perjuicio de nuestro capital ambiental natural no es ni saludable ni sostenible. Las cuentas han de llegar al fin y al cabo. El daño creciente a nuestro capital

del medio ambiente natural que resulta de los mismos procesos que producen la riqueza de nuestras sociedades debería ponernos en claro que no podemos continuar creciendo como lo hemos estado haciendo en el pasado. Una sociedad que basa su crecimiento en la degradación y agotamiento de su capital ambiental natural, no será más viable que una empresa comercial que carece de una partida adecuada de depreciación y amortización para mantener su capital y capacidad productiva. Sin embargo, esto es exactamente lo que está pasando hoy en día en mayor o menor grado en casi todos los países industrializados o en vías de desarrollo.

Los ejemplos particulares de este desenfreno son incontables en los casos de la contaminación de ríos y lagos, del aire y de las regiones costeras, en el exterminio de aves y otras formas de vida animal, y en el deterioro de áreas urbanas en los Estados Unidos y en otros países industrializados; en los países en vías de desarrollo vemos la destrucción masiva de bosques, pérdidas en gran escala de tierras productivas debidas a la erosión, desertificación y otras consecuencias de prácticas ecológicamente inaceptables en el uso de la tierra, y las pésimas condiciones de vida en las áreas urbanas "en explosión" debidas a la falta de un suministro adecuado de agua, servicio de limpia y facilidades sanitarias. También existe la amenaza creciente para los ricos tanto como para los pobres de riesgos que se dan en un ámbito más exterior, como un posible cambio climático, el daño a la capa protectora de ozono, la contaminación de cadenas alimenticias para los humanos, la contaminación de los océanos y la sobreexplotación de sus recursos vivientes.

Debe quedar claro, pues, que los patrones y prácticas actuales de crecimiento son autodestructivos e insostenibles. ¿Es entonces el alto al crecimiento la única respuesta? Permítaseme decir con todas las fuerzas de que pueda hacer acopio que el alto al crecimiento NO es la respuesta. La verdadera alternativa a este alto es un crecimiento nuevo -una visión nueva del crecimiento, tanto en las sociedades más industrializadas como en las menos desarrolladas.

El nuevo enfoque al crecimiento debe sustentarse en la eliminación del conflicto artificial y contraproducente entre ecología y economía que en el presente se encuentra incorporado en nuestro sistema de toma de decisiones económicas....

APPENDIX H Transcripts of Specialists' Summaries
of Texts 1 and 2

APPENDIX H. Transcripts of Specialists' Summaries
of Texts 1 and 2.

TEXT 1.

SPECIALIST A.

En este texto, el autor discute qué se necesita hacer si se desea lograr un equilibrio estable entre el hombre y su medio ambiente natural. Tal discusión implica lo siguiente:

1. que actualmente no hay tal equilibrio.
2. que tal equilibrio es posible.
3. que las medidas necesarias para alcanzarlo pueden ser no siempre deseadas.

En cuanto al primer punto, se dice que actualmente no hay tal equilibrio debido al incremento poblacional por un lado, y por otro, al del apetito por bienes materiales.

Sin embargo, se dice que tal equilibrio se puede alcanzar siempre y cuando se sometan a control ambos tipos de crecimiento. Si bien el crecimiento poblacional se ha nivelado en los países industrializados, en los países del tercer mundo todavía no, pero debe nivelarse. Por otra parte, es mucha la presión que se ejerce sobre el medio ambiente para satisfacer no sólo las necesidades vitales de ricos y pobres, sino además el apetito consumista insaciable de los ricos. Las presiones impuestas al medio por dicho apetito deben reducirse.

En cuanto al tercer punto, señala el autor que dado el sistema económico de los países industrializados, actualmente

les resulta difícil aceptar la necesidad de disminuir sus exigencias de recursos naturales.

Finalmente, concluye el autor que es necesario cambiar la actitud actual ante el crecimiento (poblacional y de producción) si se ha de lograr el equilibrio estable entre el hombre y su medio ambiente.

TEXT 1.

SPECIALIST B.

El autor intenta difundir su preocupación por el gran problema que nos concierne absolutamente a todos: la destrucción de los ecosistemas debida a la excesiva explotación de los recursos naturales por el hombre. Señala dos causas de esta explotación:

1. La explosión demográfica, mucho más notoria en el tercer mundo, en donde, aunque ha disminuído en cierto grado, constituye todavía una amenaza para los recursos del planeta. El autor menciona que durante las próximas tres décadas la población del tercer mundo sometera a la biósfera a fuertes presiones para satisfacer sus necesidades de subsistencia y de mejoramiento en su nivel de vida.

2. El erróneo concepto de "desarrollo económico" prevaleciente en los países industrializados, en donde "desarrollo" equivale a un aumento en la producción de bienes de consumo. Señala con alarma el autor lo intolerantes que son los países ricos a un concepto alternativo de "desarrollo". De hecho considera esta posición como una enfermedad.

Para aminorar esta extrema exigencia sobre los recursos del planeta, el autor propone:

A. El control eficaz del crecimiento de la población.

B. El cambio de actitud de los países ricos respecto al desarrollo.

Aclara que puesto que a los países subdesarrollados no se les puede negar el derecho a mejorar sus condiciones de vida, serán los ricos quienes deberán ceder en sus exigencias al planeta, si es que queremos alcanzar el equilibrio indispensable entre el hombre y el medio ambiente.

TEXT 1.

SPECIALIST C.

Dos factores se mencionan como las causas de los efectos del hombre en el medio ambiente: el crecimiento poblacional y el desarrollo económico.

El análisis divide a los países desarrollados de los subdesarrollados y a los ricos de los pobres.

Se predice que a finales de este siglo el doble de la población actual necesitará alimentos y mejoras en sus condiciones de vida. Según las predicciones, esto nos llevará a una gran crisis si no se controlan ambos crecimientos, poblacional y económico. Los efectos pueden ser el agotamiento de los recursos con el consecuente aumento en hambrunas, pestes y guerras.

Se menciona el control poblacional en países subdesarrollados como forma de minimizar estos efectos, ya que 4/5 de la población mundial estará comprendida en estos países. Los países desarrollados tendrán que cambiar su actitud hacia el crecimiento económico, donde el crecimiento material con el consecuente efecto en los recursos mundiales, es el móvil primario de estas sociedades.

Se llama la atención sobre los peligros de estas tendencias sobre un balance sostenido entre el hombre y la naturaleza, donde, de no alcanzar este balance, se corre el riesgo de una autodestrucción de nuestra sociedad.

TEXT 2.

SPECIALIST A.

En este texto el autor postula la necesidad de adoptar un nuevo enfoque de desarrollo económico, ya que el actual sistema de decisiones económicas se basa en, y propicia, un conflicto innecesario y contraproducente entre economía y ecología.

El autor señala que el crecimiento económico ha hecho posibles beneficios sin precedente para la sociedad, pero que, por depender de la degradación y agotamiento del capital del medio ambiente natural, no es ni razonable ni sostenible. (Al respecto, cita múltiples ejemplos de daños graves o irreversibles causados al medio ambiente).

Enfatiza el autor, entonces, que las consideraciones ecológicas no pueden ignorarse en nombre de las necesidades económicas. No por ignorarlas se eliminan. En todo caso, lo único que pasa es que se posponen las consecuencias, o se deja que alguien más las sufra.

Sin embargo, tampoco es la respuesta un alto al crecimiento. La alternativa planteada por el autor tanto para los países industrializados como para los países en vías de desarrollo es la de dar un nuevo enfoque al desarrollo: un crecimiento en el que el sistema de toma de decisiones económicas incorpore criterios ecológicos.

TEXT 2.

SPECIALIST B.

Muy al contrario de las posiciones ecologistas, se dice con frecuencia que la proteccion al medio ambiente es demasiado costosa y que si hay que escoger entre desarrollo y medio ambiente, el desarrollo es preferible, i.e., que la proteccion es un lujo.

El autor senala lo equivoco de este argumento en el que no se calcula la perdida que significa el dano irreparable al "capital" de bienes naturales como agua, aire, suelos, etc. Segun el autor, el desarrollo que se basa en la destruccion de este capital natural es absolutamente insostenible. No puede defenderse ni continuar ni en los paises desarrollados ni en los paises pobres.

Propone como alternativa no el alto al desarrollo (como algunos ecologistas lo han hecho), sino una nueva forma de desarrollo, en la que se logre eliminar el conflicto entre la ecologia y la economia. Propone un desarrollo economico que no tenga como premisa la destruccion del ambiente natural, sino donde exista un equilibrio entre estos dos importantes renglones de nuestra subsistencia.

TEXT 2.

SPECIALIST C.

El enfoque de los ecologistas sobre los daños al medio ambiente y la respuesta de los economistas a las alternativas son analizados.

En un principio los ecologistas se enfocaban en la limpieza o recuperación de áreas dañadas y en el control de las fuentes de contaminación. Ahora, dos extremos se han desarrollado, algunos ecologistas hacen propuestas de detener el crecimiento de sociedades industrializadas. En el otro extremo, existe la tendencia de evaluar la protección y mejoramiento del medio ambiente como un costo más que limita las ganancias económicas.

Se critica el análisis económico por dejar fuera del análisis de costo-beneficio al costo que la sociedad (como un todo) debe pagar por mantener el capital del medio ambiente natural. El crecimiento económico ha beneficiado a la sociedad, pero el desarrollo que depende de la degradación del medio ambiente no es sensato ni puede sostenerse. El pago vendrá pronto.

Ejemplos de esta destrucción son la gran cantidad de ríos, lagos, costas y aire contaminados, pérdida de vida silvestre, bosques y suelos. Las consecuencias pueden llegar a sobrepasar "barreras" ecológicas y afectar todos los sectores de la población, pobres y ricos, mediante cambios climáticos, contaminación de cadenas alimenticias y sobreexplotación de recursos.

El no crecimiento no es la alternativa, sino un "nuevo" crecimiento, un enfoque que elimine el falso conflicto entre ecología y economía, el cual es parte integral de los sistemas económicos de toma de decisiones.

APPENDIX I Instructions Used in Pilot Study
in Tests 1 and 2

Appendix I Instructions Used in Pilot Study
in Tests 1 and 2

The following is a translation into English of the instructions given to the subjects who participated in the pilot study, in Tests 1 and 2.

INSTRUCTIONS

Write a summary of the following text. Write it with the purpose of informing fellow students of yours - who might be interested in its topic, but who have not had the chance to read it - about its content.

Afterwards, include your own comments on the material read.

Write your summary and comments in Spanish.

You have a time limit of an hour to complete both tasks.

INSTRUCCIONES.

Escriba usted un resumen del siguiente texto. Escríbalo con el propósito de exponer su contenido a compañeros de clase que puedan interesarse por el tema pero que no hayan tenido la oportunidad de leer este texto.

Posteriormente, incluya también sus propios comentarios sobre el material leído.

Escriba usted su resumen y sus comentarios en español.

En total dispone usted de una hora para terminar ambas tareas.

APPENDIX J Samples of Subjects' Responses
in Tests 1 and 2 (Pilot Study)

APPENDIX J. Samples of Subjects' Responses in Tests 1 and 2.
(Pilot Study).

TEST 1.

SUBJECT 1.

El impacto del hombre en el medio ambiente resulta en un substancial crecimiento en la población mundial y en el apetito de bienes de los hombres.

Aunque el crecimiento de la población se ha estabilizado en las sociedades industrializadas y empieza a declinar en los países en vías de desarrollo, se estima que en los primeros años del siguiente siglo, la humanidad requerirá el doble de recursos. Cuatro quintas partes de la población mundial viven en países subdesarrollados y la satisfacción de sus necesidades crea significativas presiones sobre el uso de recursos naturales y la biósfera.

Las condiciones de vida de los pobres deben mejorarse y los apetitos de los ricos contenerse. El costo de no controlar el crecimiento por medio de consenso voluntario será la enfermedad y el conflicto.

De cualquier manera, el incremento de la demanda durante las tres primeras décadas del siguiente siglo será significativo y los recursos deberán acopiarse bajo las prioridades de satisfacción de las necesidades básicas de los hombres y un modo digno de vida, y de igualdad de oportunidad en la sociedad tecnológica.

Sin embargo, no hay signos de que los tiempos cambien. La hipótesis es que el crecimiento y la adquisición de bienes materiales es la base del desarrollo, es el soporte de las

políticas de los países desarrollados. Crecimiento material es la fuente del poder.

Aunque recientemente ha habido críticas a esta idea, ésta permanece como el tema dominante en la dinámica de las sociedades modernas. Se teme el no crecimiento y se postula que éste se equipara con estancamiento , con un alto costo en desempleo, seguridad social y bajos niveles de vida.

La actitud hacia el crecimiento es la base del presente dilema. Es la enfermedad que se ha desperdigado por toda la sociedad tecnológica. Esta enfermedad conlleva la capacidad de la autodestrucción de nuestra sociedad. Es necesario curarla para crear un balance substancial entre el hombre y su medio ambiente natural.

COMENTARIOS.

Es cierto que el crecimiento incontrolado tendrá un fuerte impacto en el uso de recursos y en la calidad de la vida. Controlar el crecimiento a niveles aceptables es la meta económica de todo país. El crecimiento con estabilidad es siempre deseado.

Sin embargo, existe un compromiso entre la calidad de la vida y la abundancia de recursos y nadie quiere ceder absolutamente nada. Los apetitos de los ricos son también los de los pobres.

El mejor balance de la riqueza es un bien teórico que en la dinámica de la vida real se convierte en quimera. La

actual estructura económica requiere del crecimiento y las alternativas no se conocen. Lo mejor que se puede esperar es crecimiento planeado.

Pero probablemente no sea tan grave el problema. La tesis de Adam Smith predecía que para estas épocas la humanidad debería estar agotada y no es el caso. Es prioritario dotar de calidad de vida al pobre, pero tal es el problema, no es el crecimiento per se.

El tiempo dirá.

TEST 1.

SUBJECT 7.

RESUMEN.

El documento trata uno de mis temas favoritos, y con tal de comentarlo, trataré de ser breve en mis renglones del resumen.

En terminos generales el tema tratado es el hombre y su medio (tratando de hablar del ecodesarrollo). Uno de los básicos puntos del texto es el crecimiento poblacional y su relación con el deterioro al medio. También habla del desarrollo industrial y crecimiento económico como factores que directamente son responsables del deterioro del ambiente (biósfera).

El texto trata algo de crecimiento poblacional en países industrializados y en vías de desarrollo, dando algunas proyecciones y haciendo algunas conjeturas de tipo socio-económico entre la población mundial (presente y futura). Se dan recomendaciones y se presentan observaciones (con ejemplos en algunos casos) de convivencia, desigualdad, etc. en relaciones de países ricos y pobres.

Algo notable es que el texto culpa al modelo de sociedad industrializada de su voracidad sobre la biósfera, y consecuentemente lo reprueba (al modelo) por ser responsable de las presiones sobre los recursos naturales.

Así también el sistema económico actual, como parte integral del actual modelo, obstaculiza las posibilidades inmediatas de cualquier tipo de cambio con el fin de revertir el daño causado. Así también se menciona que de seguir este patrón vigente por los bienes materiales, más y mayor será el costo ecológico por deteriorar al medio ambiente.

COMENTARIOS.

El texto me pareció muy interesante y de una gran actualidad. Creo que la manera como está expuesto (a mi gusto) requiere de algunas modificaciones, pues lo sentí muy saturado de información y a su vez ésta despeja algunas incógnitas que no son tratadas. Así también creo que le hacen falta algunos datos actualizados para reforzar ciertos mensajes, y consecuentemente haría falta hacer mención de las fuentes. Hablar no tan superficialmente de los porqués y no tan profundamente de las consecuencias, le dan a todo texto un equilibrio y en un tema como éste es importante no perderlo, por eso también considero importante incluir ciertas experiencias de organizaciones internacionales que ya vienen trabajando en ese tema, con revelaciones interesantes tanto de ámbitos regionales como mundiales. Si se responsabiliza al actual patrón de crecimiento de ser el culpable de daño al medio, creo que sería bueno mencionar a la religión, al comercio, a la educación, a las relaciones e interrelaciones mundiales,

a la evolución en general, dando quizás algunos panoramas generales del tema.

En cuanto a una hora, en particular relacionado con este texto lo sentí muy, pero muy justo.

TEST 2.

SUBJECT 2.

RESUMEN.

Los ecologistas se preocupan y hacen énfasis en detener los casos de daño al medio ambiente, como crecimiento y contaminación. Por otro lado, se dice que eso se puede hacer siempre y cuando no afecte la economía. La realidad ambiental no puede evitarse ignorándola en nombre de las necesidades económicas. El crecimiento económico ha hecho posibles muchos beneficios sin precedente, pero también nos vela la realidad del deterioro causado. El crecimiento económico que depende del agotamiento y perjuicio de nuestro capital ambiental natural, no es ni saludable ni sostenible, y esto es exactamente lo que está pasando hoy en día en mayor o menor grado en casi todos los países industrializados o en vías de desarrollo. Son incontables los ejemplos particulares de este desfreno, habiendo, como por ejemplo, contaminación de lagos, ríos, mar, etc., bosques, campos, tanto en países como E.U.A. o en países en vías de desarrollo, donde se produce erosión, desertificación y muchas consecuencias ecológicas inaceptables. Y por si fuera poco, encaramos situaciones de vida en iguales malos niveles, en las zonas urbanas, que es donde se supone que las cosas van bien.

"Quede claro pues" /los patrones y prácticas actuales de crecimiento son inaceptables! Y no es la respuesta un alto

al crecimiento, sino un nuevo enfoque al mismo, eliminando el conflicto artificial y contraproducente, entre ecología y economía, que en el presente se encuentra incorporado en nuestro sistema de toma de decisiones económicas.

COMENTARIOS.

Es interesante ver cómo se expone el hecho de que por la interminable escalada de intereses comerciales o económicos se descuide la base, lo más importante, que es el mantener un equilibrio ecológico adecuado para la supervivencia de nuestro planeta y por ende de nosotros mismos. El primer paso ya está dado, ojalá los verdaderos directos responsables lo pensarán maduramente y tomaran alguna decisión.

TEST 2.

SUBJECT 8.

RESUMEN.

El modelo de crecimiento económico que prevalece actualmente se basa en un conflicto de intereses imaginario y autodestructivo entre la ecología y la economía. Desde el punto de vista ecologista, el crecimiento económico en los países industrializados debería detenerse de una vez por todas. La opinión opuesta sostiene que la protección o mejora del medio ambiente conlleva unos costos que no nos podemos permitir, y que los intereses económicos deben prevalecer sobre los ecológicos.

Por lo que hace al segundo punto de vista, uno se pregunta si el costo para la sociedad de echar a perder el capital natural de aire, agua, suelo, fauna y flora no es mucho mayor que el capital necesario para proteger un medio ambiente del cual depende, al fin y al cabo, el bienestar y la vida de la humanidad.

Debemos, pues, afrontar la siguiente realidad: una sociedad que basa su crecimiento económico en la degradación y explotación de su medio ambiente acabará destruyéndose a sí misma, pues su bienestar depende en definitiva del bienestar mismo del medio ambiente natural. Las consecuencias de la destrucción ecológica son evidentes (contaminación de las aguas y del aire, destrucción de costas, fauna y flora,

deterioración de áreas urbanas, desaparición de bosques, erosión masiva, avance de los desiertos, etc.) y podrían conducir a consecuencias más graves: cambios climáticos, destrucción de la capa de ozono, contaminación de las cadenas alimenticias, de los mares y de su flora y fauna.

La alternativa no es un alto al crecimiento, sino la busca de un nuevo modelo de crecimiento que ponga fin a tales abusos.

COMENTARIOS.

El texto defiende la tesis de que el supuesto conflicto entre crecimiento económico e intereses ecológicos no es nada más que un conflicto imaginario. Evidentemente, mientras los defensores del modelo de crecimiento económico que actualmente prevalece en el mundo sigan insistiendo en que no nos podemos permitir los costos que supondría la mejora del medio ambiente, el mundo seguirá avanzando hacia la autodestrucción a través de la explotación desordenada de sus recursos naturales. Sin embargo, la solución no se halla tampoco en poner alto de manera radical al crecimiento económico. Lo que hace falta, tal como dice el texto, es un nuevo modelo de crecimiento que, en mi opinión, resulte en una distribución de la riqueza más equitativa: es decir, un crecimiento dirigido a los países menos desarrollados en lugar de a las necesidades menos fundamentales de los países altamente desarrollados.

APPENDIX K Data on Language Proficiency
(Subjects' ELBA Scores and Classification)

Appendix K Data on Language ProficiencySubjects' ELBA Scores and Classification

Subject (Code)	ELBA Score	Classi- ficat- ion	Subject (Code)	ELBA Score	Classi- ficat- ion	Subject (Code)	ELBA Score	Clas- sifi- cation
1	19	I	31	17	B	61	31	A
2	12	B	32	19	I	62	27	I
3	11	B	33	16	B	63	23	I
4	16	B	34	21	I	64	11	B
5	16	B	35	8	B	65	16	B
6	29	A	36	14	B	66	34	A
7	27	I	37	28	A	67	69	A
8	14	B	38	14	B	68	30	A
9	18	I	39	34	A	69	13	B
10	7	B	40	24	I	70	14	B
11	18	B	41	23	I	71	11	B
12	16	B	42	29	A	72	20	I
13	17	B	43	45	A	73	31	A
14	21	I	44	24	I	74	31	A
15	30	A	45	16	B	75	37	A
16	10	B	46	14	B	76	29	A
17	24	I	47	23	I	77	13	B
18	19	I	48	11	B	78	14	B
19	31	A	49	55	A	79	20	I
20	24	I	50	6	B	80	24	I
21	15	B	51	21	I	81	22	I
22	15	B	52	37	A	82	34	A
23	26	I	53	32	A	83	19	I
24	36	A	54	23	I	84	19	I
25	19	I	55	56	A	85	41	A
26	52	A	56	17	B	86	21	I
27	29	A	57	19	I	87	17	B
28	32	A	58	29	A	88	40	A
29	23	I	59	33	A	89	25	I
30	30	A	60	24	I	90	42	A

Where:

B = Beginners

I = Intermediate

A = Advanced

APPENDIX L Data on Reading Comprehension
(Subjects' Scores in Tests 1 and 2)

Appendix L Data on Reading Comprehension
(Subjects' Scores on Tests 1 and 2)

<u>TEST 1</u>						
I	II	III	IV	V	VI	VII
SUBJECT	LANGUAGE OF TEST*	JUDGE A	JUDGE B	DIFFERENCE FROM A TO B	HIGHEST OF A OR B	AVERAGE OF A AND B
1	S	2.5	2	.5 -	2.5	2.25
2	S	3.5	2	1.5 -	3.5	2.75
3	E	1	1.5	.5 +	1.5	1.25
4	E	1	1	0	1	1
5	E	2.5	2	.5 -	2.5	2.25
6	S	2.5	2	.5 -	2.5	2.25
7	E	2.5	2	.5 -	2.5	2.25
8	E	1.5	1.5	0	1.5	1.5
9	S	5	5	0	5	5
10	S	3.5	2.5	1 -	3.5	3
11	E	2.5	1	1.5 -	2.5	1.75
12	S	3.5	3	.5 -	3.5	3.25
13	E	2.5	1.5	1 -	2.5	2
14	S	4	5	1 +	5	4.5
15	E	3.5	5	1.5 +	5	4.25
16	S	3.5	2	1.5 -	3.5	2.75
17	E	2.5	1.5	1 -	2.5	2
18	S	4.5	3	1.5 -	4.5	3.75
19	E	2.5	2.5	0	2.5	2.5
20	E	3	2	1 -	3	2.5
21	E	1	2	1 +	2	1.5
22	S	2.5	2	.5 -	2.5	2.25
23	S	3.5	3	.5 -	3.5	3.25
24	E	4.5	3	1.5 -	4.5	3.75
25	S	4.5	4.5	0	4.5	4.5
26	S	4	4	0	4	4
27	E	3.5	3.5	0	3.5	3.5
28	S	3.5	5	1.5 +	5	4.25
29	E	3.5	3.5	0	3.5	3.5
30	E	3.5	3.5	0	3.5	3.5

*Where:

E = English

S = Spanish

TEST 1

I	II	III	IV	V	VI	VII
SUBJECT	LANGUAGE OF TEST	JUDGE A	JUDGE B	DIFFERENCE FROM A TO B	HIGHEST OF A OR B	AVERAGE OF A AND B
31	S	2.5	4	1.5 +	4	3.25
32	S	4.5	3.5	1 -	4.5	4
33	S	3.5	3	0.5 -	3.5	3.25
34	S	2.5	2.5	0	2.5	2.5
35	S	2.5	2.5	0	2.5	2.5
36	S	3.5	3	.5 -	3.5	3.25
37	E	3.5	2.5	1 -	3.5	3
38	E	2.5	1.5	1 -	2.5	2
39	S	5	3.5	1.5 -	5	4.25
40	S	3.5	3.5	0	3.5	3.5
41	E	4	4.5	.5 +	4.5	4.25
42	S	3	4.5	1.5 +	4.5	3.75
43	E	4.5	4.5	0	4.5	4.5
44	E	3.5	2	1.5 -	3.5	2.75
45	S	3.5	2.5	1 -	3.5	3
46	S	3.5	3	.5 -	3.5	3.25
47	E	4	3.5	.5 -	4	3.75
48	E	2.5	1	1.5 -	2.5	1.75
49	E	3.5	5	1.5 +	5	4.25
50	E	2	2	0	2	2
51	S	3.5	4	.5 +	4	3.75
52	E	3.5	2	1.5 -	3.5	2.75
53	S	4	4.5	.5 +	4.5	4.25
54	S	4	4	0	4	4
55	E	5	5	0	5	5
56	E	2.5	1	1.5 -	2.5	1.75
57	S	5	4.5	.5 -	5	4.75
58	S	3	3	0	3	3
59	S	3	2	1 -	3	2.5
60	E	4	3.5	.5 -	4	3.75

TEST 1

I	II	III	IV	V	VI	VII
SUBJECT	LANGUAGE OF TEST	JUDGE A	JUDGE B	DIFFERENCE FROM A TO B	HIGHEST OF A OR B	AVERAGE OF A AND B
61	E	4	4.5	.5 +	4.5	4.25
62	S	2.5	3	.5 +	3	2.75
63	S	2.5	2	.5 -	2.5	2.25
64	S	3.5	3.5	0	3.5	3.5
65	S	4	3	1 -	4	3.5
66	E	2.5	3	.5 +	3	2.75
67	S	5	4.5	.5 -	5	4.75
68	S	4	4.5	.5 +	4.5	4.25
69	S	3	3	0	3	3
70	E	2.5	1	1.5 -	2.5	1.75
71	E	3.5	2	1.5 -	3.5	2.75
72	E	3.5	1.5	2 -	3.5	2.5
73	S	4	4	0	4	4
74	S	2.5	2.5	0	2.5	2.5
75	E	3.5	4	.5 +	4	3.75
76	S	3	2	1 -	3	2.5
77	E	2.5	2.5	0	2.5	2.5
78	S	3.5	2.5	1 -	3.5	3
79	E	2.5	2.5	0	2.5	2.5
80	E	2.5	1.5	1 -	2.5	2
81	E	3.5	3.5	0	3.5	3.5
82	E	3.5	3	.5 -	3.5	3.25
83	E	3.5	2	1.5 -	3.5	2.75
84	S	2.5	1.5	1 -	2.5	2
85	S	4	3.5	.5 -	4	3.75
86	E	4	3	1 -	4	3.5
87	E	4	2	2 -	4	3
88	S	3	3	0	3	3
89	E	3	2	1 -	3	2.5
90	E	3.5	2	1.5 -	3.5	2.75

TEST 2

I	II	III	IV	V	VI	VII
SUBJECT	LANGUAGE OF TEST	JUDGE A	JUDGE B	DIFFERENCE FROM A TO B	HIGHEST OF A OR B	AVERAGE OF A AND B
1	E	2.5	2.5	0	2.5	2.5
2	E	1	2.5	1.5 +	2.5	1.75
3	S	4	4	0	4	4
4	S	2	2	0	2	2
5	S	3.5	2.5	1 -	3.5	3
6	E	3	4	1 +	4	3.5
7	S	1	2	1 +	2	1.5
8	S	3.5	3.5	0	3.5	3.5
9	E	4	4	0	4	4
10	E	1.5	3	1.5 +	3	2.25
11	S	2.5	4.5	2 +	4.5	3.5
12	E	3	3	0	3	3
13	S	2.5	3.5	1 +	3.5	3
14	E	4	3.5	.5 -	4	3.75
15	S	4	4.5	.5 +	4.5	4.25
16	E	1	1	0	1	1
17	S	4.5	3	1.5 -	4.5	3.75
18	E	4	3	1 -	4	3.5
19	S	4.5	4.5	0	4.5	4.5
20	S	3.5	5	1.5 +	5	4.25
21	S	4.5	4	.5 -	4.5	4.25
22	E	2.5	3	.5 +	3	2.75
23	E	5	5	0	5	5
24	S	5	5	0	5	5
25	E	4.5	5	.5 +	5	4.75
26	E	4	4	0	4	4
27	S	3.5	4.5	1 +	4.5	4
28	E	3.5	4	.5 +	4	3.75
29	S	4	4	0	4	4
30	S	4	4	0	4	4

TEST 2

I	II	III	IV	V	VI	VII
SUBJECT	LANGUAGE OF TEST	JUDGE A	JUDGE B	DIFFERENCE FROM A TO B	HIGHEST OF A OR B	AVERAGE OF A AND B
31	E	1	1.5	.5 +	1.5	1.25
32	E	1	1	0	1	1
33	E	3.5	4	.5 +	4	3.75
34	E	1.5	3	1.5 +	3	3.25
35	E	2	2	0	2	2
36	E	1.5	3	1.5 +	3	2.25
37	S	3.5	3.5	0	3.5	3.5
38	S	5	3.5	1.5 -	5	4.5
39	E	5	4.5	.5 -	5	4.75
40	E	5	4.5	.5 -	5	4.75
41	S	3.5	4	.5 +	4	3.75
42	E	2	2.5	.5 +	2.5	2.25
43	S	5	5	0	5	5
44	S	3.5	4.5	1 +	4.5	4
45	E	2.5	1	1.5 -	2.5	1.75
46	E	5	4.5	.5 -	5	4.75
47	S	5	3.5	1.5 -	5	4.25
48	S	3	2.5	.5 -	3	2.75
49	S	3.5	4	.5 +	4	3.75
50	S	5	5	0	5	5
51	E	2.5	2.5	0	2.5	2.5
52	S	4	3.5	.5 -	4	3.75
53	E	3.5	4	.5 +	4	3.75
54	E	4	3	1 -	4	3.5
55	S	4	3.5	.5 -	4	3.75
56	S	4	3	1 -	4	3.5
57	E	2	2	0	2	2
58	E	4	3.5	.5 -	4	3.75
59	E	4.5	3	1.5 -	4.5	3.75
60	S	5	5	0	5	5

TEST 2

I	II	III	IV	V	VI	VII
SUBJECT	LANGUAGE OF TEST	JUDGE A	JUDGE B	DIFFERENCE FROM A TO B	HIGHEST OF A OR B	AVERAGE OF A AND B
61	S	5	5	0	5	5
62	E	2.5	3	.5 +	3	2.75
63	E	3.5	3.5	0	3.5	3.5
64	E	5	4.5	.5 -	5	4.75
65	E	2	2	0	2	2
66	S	2	3	1 +	3	2.5
67	E	4.5	3	1.5 -	4.5	3.75
68	E	4	3.5	.5 -	4	3.75
69	E	2.5	2.5	0	2.5	2.5
70	S	3.5	2.5	1 -	3.5	3
71	S	4.5	3.5	1 -	4.5	4
72	S	4	3.5	.5 -	4	3.75
73	E	4	3.5	.5 -	4	3.75
74	E	1.5	2.5	1 +	2.5	2
75	S	4	3.5	.5 -	4	3.75
76	E	4.5	3.5	1 -	4.5	4
77	S	3	2	1 -	3	2.5
78	E	3	2	1 -	3	2.5
79	S	3.5	4	.5 +	4	3.75
80	S	3.5	3	.5 -	3.5	3.25
81	S	3	3	0	3	3
82	S	3.5	2.5	1 -	3.5	3
83	S	2.5	3	.5 +	3	2.75
84	E	1	2	1 +	2	1.5
85	E	4	3	1 -	4	3.5
86	S	5	3.5	1.5 -	5	4.25
87	S	5	5	0	5	5
88	E	4	3	1 -	4	3.5
89	S	4.5	3.5	1 -	4.5	4
90	S	5	3.5	1.5 -	5	4.25

APPENDIX M Tables Used in the Preparation of Main Table of
Analysis of Variance (Complete ANOVA)

Appendix M Tables Used in the Preparation of the Main Table of
Analysis of Variance (Complete ANOVA)

TABLE M.1

Original data. SOLO scores: a three-way classification table, according to Language of Test (English or Spanish), Order of Administration of Test (First or Second), and Level of Proficiency in English (Beginners, Intermediate or Advanced).

	English		Spanish	
	First*	Second**	First**	Second*
B e g i n n e r s	3	4.75	3.5	5
	2.75	2	3.5	4
	2.5	4.75	3.25	2.5
	2.25	3.75	3.25	3
	2	3	3.25	5
	2	2.25	3.25	4.5
	2	1.25	3.25	3
	1.75	2.5	3	3.5
	1.75	2.5	3	3.5
	1.75	2.25	3	3
	1.75	1.75	3	2.75
	1.5	1.75	2.75	4.25
	1.5	1	2.75	3.5
	1.25	2	2.5	4
	1	2.75	2.25	2
I n t e r m e d i a t e	4.25	4	5	3.75
	3.75	2	4.75	5
	3.75	4.75	4.5	4.25
	3.5	3.75	4.5	4.25
	3.5	3.5	4	4
	3.5	1	4	3
	2.75	3.5	3.75	4
	2.75	2.5	3.75	2.75
	2.5	4.75	3.5	4.25
	2.5	5	3.25	4
	2.5	2.75	2.75	3.75
	2.5	2.25	2.5	3.75
	2.25	3.5	2.25	1.5
	2	2.5	2.25	3.75
	2	1.5	2	3.25

A d v a n c e d	5	3.75	4.75	3.75
	4.5	4.75	4.25	5
	4.25	3.75	4.25	5
	4.25	3.75	4.25	4.25
	4.25	4	4	3.75
	3.75	3.75	4	5
	3.75	3.75	3.75	3.75
	3.5	3.5	3.75	4
	3.5	2.25	3.75	4
	3.25	3.75	3	3
	3	3.5	3	3.5
	2.75	4	2.5	4.25
	2.75	3.75	2.5	3.75
	2.75	2	2.5	2.5
	2.5	3.5	2.25	4.5

TABLE M.2

Obtained by totalling the repeated values ("replications") within each treatment combination. (From Table M.1).

	English		Spanish	
	First	Second	First	Second
Beginners	28.75	38.25	45.5	53.5
Intermediate	44	47.25	52.75	55.25
Advanced	53.75	53.75	52.5	60

N.B. Each entry in the above table is the sum of 15 scores.

TABLE M.3.1

Obtained from Table M.2 by summing over Level of Proficiency

English		Spanish	
First	Second	First	Second
126.5	139.25	150.75	168.75

N.B. Each entry in the above table is the sum of 45 scores.

TABLE M.3.2

Obtained from Table M2 by summing over Language of Test

First			Second		
Beginners	Intermediate	Advanced	Beginners	Intermediate	Advanced
74.25	96.75	106.25	91.75	102.05	113.75

N.B. Each entry in the above table is the sum of 30 scores.

TABLE M.3.3

Obtained from Table M2 by summing over Order of Administration

English			Spanish		
Beginners	Intermediate	Advanced	Beginners	Intermediate	Advanced
67	91.25	107.5	99	108	112.5

N.B. each entry in the above table is the sum of 30 scores.

TABLE M.4.1

Obtained by summation from Tables M.3.1 or M.3.2

First	Second
277.25	308

Grand Total = T = 585.25
for N = 180 items

N.B. Each entry is the sum of 90 scores.

TABLE M.4.2

Obtained by summation from Tables M.3.2 or M.3.3

Beginners	Intermediate	Advanced
166	199.25	220

Grand Total = T = 585.25
for N = 180 items

N.B. Each entry is the sum of 60 scores.

TABLE M.4.3

Obtained by summation from Tables M.3.1 or M.3.3

English	Spanish
265.75	319.5

Grand Total = T = 585.25
for N = 180 items

N.B. Each entry is the sum of 90 scores.

$$\begin{aligned} \text{Correction Factor} &= \frac{T^2}{N} \\ &= \frac{585.25^2}{180} = 1902.8753 \end{aligned}$$

TABLE M.5

Obtained by summation from table M.2

	English First + Spanish Second	Spanish First + English Second
Beginners	82.25	83.75
Intermediate	99.25	100
Advanced	113.75	106.25

Grand Total = 585.25
for N = 90 items
(15 per cell)

$$\begin{aligned} \text{Correction Factor} &= \frac{T^2}{N} \\ &= \frac{585^2}{90} \\ &= 3805.7506 \end{aligned}$$

N.B. Each entry is the sum of 30 scores, obtained by 15 subjects, and which are therefore taken as 15 items per cell (i.e., 15 repeated measures).

APPENDIX N Tables Used in the Preparation of the Tables of
Analysis of Variance by "Language" (of Test),
and by "Level" (of Proficiency). (Breakdown
Analyses of Variance)

Appendix N Tables Used in the Preparation of the Tables of
Analysis of Variance by "Language" (of Test),
and by "Level" (of Proficiency). (Breakdown
Analyses of Variance).

Breakdown Analysis of Variance by "Language" (of Test)

ENGLISH:

Table N.1 (English)

(from Table M.2)

	English First	English Second
B	28.75	38.25
I	44	47.25
A	53.75	53.75

Each entry is the sum of 15 scores.

TABLE N.2 (English)

B	I	A
67	91.25	107.5

Each entry is the sum of 30 scores.

TABLE N.3 (English)

English First	English Second
126.5	139.25

Each entry is the sum of 45 scores

$$\begin{aligned}
 \text{Grand Total} &= 265.75 \\
 \text{for } N &= 90 \text{ items} \\
 \text{Correction Factor} &= \frac{T^2}{N} \\
 &= \frac{265.75^2}{90} = 784.7
 \end{aligned}$$

Breakdown Analysis of Variance by "Language" (of Test)

SPANISH:

TABLE N.1 (Spanish)

(from Table M.2)

	Spanish First	Spanish Second
B	45.5	53.5
I	52.75	55.25
A	52.5	60

Each entry is the sum of 15 scores.

TABLE N.2 (Spanish)

B	I	A
99	108	112.5

Each entry is the sum of 30 scores.

TABLE N.3 (Spanish)

Spanish First	Spanish Second
150.75	168.75

Each entry is the sum of 45 scores.

$$\begin{aligned}
 \text{Grand Total} &= 319.5 \\
 \text{for } N &= 90 \text{ items} \\
 \text{Correction Factor} &= \frac{319.5^2}{90} \\
 &= 1134.225
 \end{aligned}$$

Breakdown Analysis of Variance by "Level" (of Proficiency)

ADVANCED:

TABLE N.1 (Advanced)

	First	Second
English	53.75	53.75
Spanish	52.5	60

Each entry is the sum of 15 scores.

TABLE N.2 (Advanced)

First	Second
106.25	113.75

Each entry is the sum of 30 scores.

TABLE N.3 (Advanced)

English	Spanish
107.5	112.5

Each entry is the sum of 30 scores.

$$\begin{aligned}
 \text{Grand Total} &= 220 \\
 \text{for } N &= 60 \text{ items} \\
 \text{Correction Factor} &= \frac{220^2}{60} \\
 &= 806.6666
 \end{aligned}$$

Breakdown Analysis of Variance by "Level" (of Proficiency)

INTERMEDIATE:

Table N.1 (Intermediate)

	First	Second
English	44	47.25
Spanish	52.75	55.25

Each entry is the sum of 15 scores.

TABLE N.2 (Intermediate)

First	Second
96.75	102.5

Each entry is the sum of 30 scores.

TABLE N.3 (Intermediate)

English	Spanish
91.25	108

Each entry is the sum of 30 scores.

$$\begin{aligned}
 \text{Grand Total} &= \frac{199.25}{\text{for } N = 60 \text{ items}} \\
 \text{Correction Factor: } &\frac{199.25^2}{60} \\
 &= 661.676
 \end{aligned}$$

Breakdown Analysis of Variance by "Level" (of Proficiency)

BEGINNERS:

Table N.1 (Beginners)

	First	Second
English	28.75	38.25
Spanish	45.5	53.5

Each entry is the sum of 15 scores.

TABLE N.2 (Beginners)

First	Second
74.25	91.75

Each entry is the sum of 30 scores.

TABLE N.3 (Beginners)

English	Spanish
67	99

Each entry is the sum of 30 scores.

$$\begin{aligned}
 \text{Grand Total} &= 166 \\
 \text{for } N &= 60 \text{ items} \\
 \text{Correction Factor} &= \frac{166^2}{60} \\
 &= 459.2666
 \end{aligned}$$

APPENDIX O Transcripts of Subjects' Responses
Representative of Each Level of the
SOLO Taxonomy

Appendix O. Transcripts of Subjects' Responses Representative of Each Level of the SOLO Taxonomy.

SUBJECT: 16.

SOLO LEVEL: Prestructural.

LANGUAGE OF TEST: English.

LEVEL OF PROFICIENCY: Beginner.

ORDER OF ADMINISTRATION: Second.

SUMMARY:

La tecnología moderna ha tenido o más bien, ha sido un impacto fuerte en los EE.UU.

Esto ha traído grandes polémicas, y se ha presentado un conflicto entre la economía y la "environmental factors".

La economía ha sufrido algunos cambios en sus cálculos y sus costos. Mientras que la sociedad ha tenido que mantener el capital natural "environmental", del agua, aire, sol y la vida vegetal y animal así como la vida del hombre. ¿Pero de quién depende esto? Se preguntan.

No es necesario preservar la situación. Se requiere reconsiderar los costos, para que la economía se beneficie.

Se sabe que la economía depende del pueblo, y por lo tanto deben contribuir a los modos y medios de producción.

Sólo así el sistema económico dará solución a las demandas del país.

COMMENTS:

Se me hizo bastante difícil de entender, pues por lo regular cuando uno lee algun texto en inglés más o menos se da cuenta de lo que trata por los "cognados".

Sin embargo, aquí son pocos, y la palabra clave creo que era "environmental", y no sé su significado, por lo que se me dificultó la lectura.

SUBJECT: 8.

SOLO LEVEL: Transitional (from Prestructural to Unistructural).

LANGUAGE OF TEST: English.

LEVEL OF PROFICIENCY: Beginner.

ORDER OF ADMINISTRATION: First.

SUMMARY:

La población crece en las sociedades industrializadas, y para el próximo siglo la tierra tendrá el doble de la población actual y las reservas naturales no serán suficientes más que para una minoría, pues son improbables las condiciones de vida.

Por lo tanto, hay que establecer métodos voluntarios de control de población y evitar consecuencias trágicas o conflictos.

COMMENTS:

En lo personal aún soy muy insegura en el idioma inglés, pues comprendí frases sueltas y no estoy segura de haber puesto lo correcto pues son muy dispersas.

Hay palabras clave que no sé su significado, por lo que me costó aún más trabajo entender el texto.

SUBJECT: 65.

SOLO LEVEL: Unistructural.

LANGUAGE OF TEST: English.

LEVEL OF PROFICIENCY: Beginner.

ORDER OF ADMINISTRATION: Second.

SUMMARY:

Las sociedades industrializadas han hecho posibles muchos beneficios para la humanidad, pero al mismo tiempo esta industrialización ha provocado transtornos ecológicos. Se ha contaminado el mar, se ha causado la muerte de muchos animales, se han deteriorado algunas zonas, han habido grandes destrucciones de bosques, erosiones, transformaciones en el clima, contaminación en los mares.

COMMENTS:

Yo creo que esto sí es un gran problema que puede exterminar al mundo (a todos los seres humanos) por completo. Quizás no sea muy rápida la destrucción, pero ésta se va a dar a través del tiempo.

Es triste que poco a poco ese afán de producción e industrialización termine con el hombre.

SUBJECT: 70.

SOLO LEVEL: Transitional (from Uniststructural to Multi-
structural).

LANGUAGE OF TEST: English.

LEVEL OF PROFICIENCY: Beginner.

ORDER OF ADMINISTRATION: First.

SUMMARY:

Esta lectura nos habla del crecimiento de la población dentro de las sociedades industrializadas en cuanto a sus condiciones de vida, nos explica que se han tomado algunos métodos para controlar el trágico incremento de población voluntariamente.

En algunos países han tenido que reducir sus demandas económicas, y ahora optarán por el no-crecimiento, ya que el crecimiento de la población y la demanda económica llevarán a la destrucción de la sociedad.

COMMENTS:

El tema que toca la lectura me parece muy importante y cierto, ya que si no tomamos medidas se llegará a la destrucción de nosotros mismos.

Me costó un poco de trabajo entenderlo, pero por palabras claves pude sacar una idea del texto.

SUBJECT: 62.

SOLO LEVEL: Multistructural.

LANGUAGE OF TEST: Spanish.

LEVEL OF PROFICIENCY: Intermediate.

ORDER OF ADMINISTRATION: First.

SUMMARY:

En las sociedades industrializadas el crecimiento de la población ya se ha nivelado, mientras que en las sociedades en vías de desarrollo continúa siendo aún un problema severo. Esto es un conflicto grave, ya que para los primeros años del próximo siglo el planeta tendrá el doble de la población actual. Por otra parte, los pobres, que constituyen la mayor parte de la población, resultarán ser los más afectados, por lo que, tarde o temprano algo o alguien tendrá que ceder o cambiar.

Aunque algunos optimistas opinan que hay control en los índices de crecimiento, persisten los problemas como el de alimentar a millones de bocas, enfrentar enfermedades, etc., para tener una vida digna, satisfaciendo las necesidades básicas.

Los países desarrollados no intentan modificar sus sistemas económicos, ya que para ellos el bienestar de las sociedades puede asegurarse con el crecimiento continuo en

el sentido material, a pesar de la controversia que se ha suscitado.

Este es uno de los problemas más graves, que nos debe importar a todos haciéndole frente.

COMMENTS:

La información es bastante clara, e interesante, pudiendo así elaborar un resumen.

Como el mismo texto indica, es un problema que nos atañe a todos, y que todos debemos cooperar para tener una mejor relación con el medio ambiente, sin el cuál no seríamos quienes somos, o simplemente no existiríamos.

SUBJECT: 81.

SOLO LEVEL: Transitional (from Multistructural to Relational).

LANGUAGE OF TEST: English.

LEVEL OF PROFICIENCY: Intermediate.

ORDER OF ADMINISTRATION: First.

SUMMARY:

El hombre con su crecimiento, aumento de su población, y aumento en la cantidad de alimentos que necesita, ha causado un enorme impacto sobre los recursos naturales.

El crecimiento en sociedades industrializadas ha descendido, pero en partes del mundo sigue siendo alto. Es posible que la población se duplique para el siguiente siglo; y esto crea una gran presión sobre los recursos naturales y la biósfera.

Aparte del aumento del número de bocas que alimentar, están la necesidad de mejores condiciones de vida y de oportunidades iguales para gozar de los beneficios de la tecnología. Esto puede alcanzarse o mejorar si las sociedades industrializadas reducen las presiones que ocasionan sobre la biósfera y los recursos naturales.

La gente de algunos países desarrollados ha creado campañas en contra de la sobrepoblación, porque ocasiona un retraso en el crecimiento económico, lo que ocasiona una baja

en el modo de vida.

Por todo lo antes dicho es necesario un justo balance entre el hombre y los recursos naturales.

COMMENTS:

El texto marca uno de los graves problemas de nuestros días, que es la sobrepoblación y el grave daño que ocasiona sobre los recursos naturales, que son renovables, pero no crecen al mismo ritmo con que crece la sociedad. Pienso que se debe tomar conciencia sobre este problema, tratar de resolverlo.

En cuanto a la forma del escrito, hubo partes que resultaron un tanto obscuras en su significado, pero el conocimiento de algunas palabras hacen del texto un tanto comprensible.

SUBJECT: 26.

SOLO LEVEL: Relational.

LANGUAGE OF TEST: Spanish.

LEVEL OF PROFICIENCY: Advanced.

ORDER OF ADMINISTRATION: First.

SUMMARY:

El presente texto describe cuáles son los efectos del impacto del hombre en el medio ambiente; en cuanto al crecimiento de la población humana y al crecimiento de consumo de bienes materiales. Asimismo, presenta la idea de controlar estos fenómenos.

Por un lado, señala que el aumento de la población en países industrializados ha sido ya nivelada y en los países en vías de desarrollo, a pesar de no estar bajo control total, parece ir disminuyendo.

Sin embargo, por otro lado, apunta el hecho de que en los primeros años del siglo XXI la población será el doble de la que existe en la actualidad. La gravedad de esto estriba en los problemas que se desencadenarán como por ejemplo: presiones en los recursos naturales y biósfera, demandas de las clases bajas para mejorar sus condiciones socioeconómicas y el aumento de la concentración de la riqueza en pocas manos.

Algunas de las medidas que menciona este artículo para frenar la explosión demográfica son: métodos de control de

la población voluntarios o por las consecuencias trágicas y traumáticas del hambre y la enfermedad, etc.

No obstante, no se puede negar a la población y sobre todo a los sectores más pobres el derecho a alcanzar mejores niveles de vida y de gozar de los beneficios que los avances tecnológicos de la sociedad ofrece. Esto puede suceder sólo si los países industrializados reducen las presiones que ejercen sobre la biósfera y los recursos naturales.

Pero resulta muy poco probable que éstos accedan a reducir voluntariamente sus demandas. Debido a que perderían toda la riqueza y poderío del que gozan actualmente a costa de los países subdesarrollados.

Finalmente, es esta actitud ante el crecimiento la que lleva consigo el potencial para la autodestrucción de nuestra sociedad y lo que precisamente debemos evitar.

COMMENTS:

Me parece que es un texto interesante, el cual presenta una problemática (explosión demográfica y explotación de países industrializados a países "en vías de desarrollo"), la cuál desde hace varios años ha comenzado a preocupar y ante la cual se deben de plantear estrategias para resolverla.

SUBJECT: 67.

SOLO LEVEL: Transitional (from Relational to Extended Abstract).

LANGUAGE OF TEST: English.

LEVEL OF PROFICIENCY: Advanced.

ORDER OF ADMINISTRATION: Second.

SUMMARY:

Este texto continúa de cierta manera con lo leído ya anteriormente. En esta ocasión, se dedica a analizar el conflicto existente entre lo que podríamos llamar las decisiones económicas para el desarrollo contra los costos ecológicos y ambientales de dicho desarrollo.

Se nos presentan opiniones acerca de los planes para disminuir la contaminación de los recursos o en otros casos su agotamiento. Según el autor, no podemos sustraer de los costos del desarrollo industrial los propios costos que se generan por la degradación del medio ambiente en el planeta: "No es posible evitar (ocultar) las realidades ecológicas, simplemente ignorándolas en nombre de las necesidades económicas". Pues bien, ¿Es suficiente con pedir la opción de "crecimiento cero" tanto para las potencias industrializadas como para los países en vías de desarrollo? El autor nos dice que no, que es necesaria una nueva estrategia de crecimiento económico y social que vaya de acuerdo con el buen uso, aprovechamiento y conservación de los recursos

naturales. Es decir, que las necesidades económicas no vayan en contradicción con los costos ambientales ya que deben ser tomados muy en cuenta por sus consecuencias a largo plazo (ahora ya no tan largo).

COMMENTS:

El texto me parece muy acertado en cuanto a los aspectos del descuido de los recursos naturales por ser su costo de mantenimiento muy elevado. Pero es cierto también que los costos a largo plazo serán aún mayores, e inclusive irreversibles.

También los aspectos en cuanto al crecimiento excesivo de las ciudades, sobre todo en los países del Tercer Mundo, es otra situación que debe preocuparnos a todos.

Creo que el texto es muy adecuado para ver la perspectiva económica y social del problema de los recursos naturales.

SUBJECT: 43.

SOLO LEVEL: Extended Abstract.

LANGUAGE OF TEST: Spanish.

LEVEL OF PROFICIENCY: Advanced.

ORDER OF ADMINISTRATION: Second.

SUMMARY:

Los diferentes puntos de vista entre las tendencias ecologistas y las prácticas económicas actuales han dado como resultado que no se lleve a cabo una acción eficaz en contra de las fuentes de contaminación, tanto las actuales como las que podrían surgir con el desarrollo de las industrias, ciudades, etc.

Los industriales aplican acciones ecologistas siempre y cuando éstas estén en un rango económico, cuando las acciones pasan ese rango, los cálculos de costo-beneficio hacen que recorten las acciones, para que éstas entren en el presupuesto. Estas acciones de recorte, o simplemente de no aplicación de las medidas que son necesarias para preservar el medio ambiente, son totalmente infundadas, ya que el preservar el medio ambiente no se puede ver como un costo, no se deben ver frenadas las medidas por causas de índole económico.

El crecimiento industrial ha hecho que la mayoría de las acciones que tienden a proteger el ambiente sean vistas

como un costo de los artículos que se producen, pero supeditadas a la misma industria. Se puede ver que en un futuro se tiene que tener otra visión: de una productividad que no dañe a la naturaleza y que esté supeditada a la naturaleza misma, ya que es lo más importante que se tiene, ya que al sufrir un daño irreparable, el daño al final de cuentas recaería en nosotros.

No se debe crecer a costa de contaminar lagos, ríos, la atmósfera, etc. Se podría pensar que la solución es no crecer; pero la solución más viable es crecer de un modo diferente, esto es, no basar el crecimiento en la degradación de nuestro medio ambiente, en el exterminio de especies, contaminación y aniquilamiento de bosques, etc., ya que esto podría cambiar el clima, contaminar las cadenas alimenticias, etc.

El crecimiento en un futuro debe estar equilibrado con el medio ambiente.

COMMENTS:

La contaminación ambiental es un problema que se tiene que resolver. Estoy de acuerdo con que no se deben poner restricciones de tipo económico a las medidas que tienden a proteger el medio ambiente.

Se debe pensar de una manera diferente respecto al crecimiento. Debemos pensar que los artículos que consumimos

son en sí un desgaste al medio ambiente, por lo que se debe de ver qué es lo que realmente se necesita para vivir bien, y no caer en un consumismo que crea muchos artículos que no son indispensables para una buena vida, pero sí son esenciales para una economía que necesita crecer a cada momento.

El poner los intereses industriales en un segundo plano, exige un cambio del sistema económico actual. Este cambio creo sería benéfico, aunque actualmente lo creo poco factible. Lo que se debe hacer es una conciencia de los consumidores y de ahí tratar de cambiar un poco el mercado y con esto la producción de ciertos artículos que dañan el medio ambiente. Ya que una acción a gran escala en contra de la contaminación es casi imposible actualmente.

BIBLIOGRAPHY

- Alber-DeWolf, R. 1984. "Term-formation processes in English, German, French and Russian." In Pugh, A.K., and Ulijn, J.M. (Eds.). 1984.
- Alderson, J.C. 1978. The Production of Pedagogic Materials at UNAM. Research and Development Unit, Report No.12, mimeo, CELE, UNAM, Mexico.
- Alderson, J.C. 1984. "Reading in a foreign language: a reading problem or a language problem?" In Alderson, J.C., and Urquhart, A.H. (Eds.). 1984.
- Alderson, J.C., and Alvarez, G. 1977. The Development of Strategies for the Assignment of Semantic Information to Unknown Lexemes in Text. Research and Development Unit, Report No.10, mimeo, CELE, UNAM, Mexico.
- Alderson, J.C., and Richards, S. 1977. Difficulties Which Students Encounter When Reading Texts in English. Research and Development Unit, Report No. 8, mimeo, CELE, UNAM, Mexico.
- Alderson J.C., and Urquhart, A.H. (Eds.). 1984. Reading in a Foreign Language. New York: Longman.
- Alderson, J.C., and Williamson, M. 1978. A Framework for the Evaluation of Teaching Materials. Research and Development Unit, Report No.13, mimeo, CELE, UNAM, Mexico.
- Alderson J.C., Bastien, S., and Madrazo, A.M. 1977. A Comparison of Reading Comprehension in English and Spanish. Research and Development Unit Report No.9, mimeo, CELE, UNAM, Mexico.
- Alderson, J.C., Bastien, S., and Richards, S. 1978. A Framework for the Production of ESP Courses. Research and Development Unit, Report No.11, mimeo, CELE, UNAM, Mexico.
- Allen, J.P.B., and Davies, A. (Eds.). 1977. Testing and Experimental Methods. Vol.4 of the Edinburgh Course in Applied Linguistics. London: Oxford University Press.
- Allen, J.P.B., and Widdowson, H.G. 1974, "Teaching the Communicative Use of English." International Review of Applied Linguistics in Language Teaching, 1974, pp. 1-21, cited in Coulthard, M. 1977 , p.148.
- Alvarez, G. 1980. "Factores Sociales, Económicos y Políticos que Influyen en la Enseñanza de Lenguas Extranjeras." Comunicación, la Revista del Profesor de Idiomas. Vol.4, No. 1, April, 1980, Lima, Peru.
- Alvarez, G., Chasan, M., and Galicia, F. 1980. Consideraciones sobre los Logros y las Perspectivas de los Cursos con Objetivos Específicos en el CELE. Xalapa, Mexico: ESP Regional Symposium, (typescript).

- Alvarez, G., and Mackay, R. 1976. English for Science and Technology: The Handling of Vocabulary in Reading Comprehension Text Books. Research and Development Unit, Report No.5, mimeo, CELE, UNAM, Mexico.
- American Psychological Association. 1966. Standards for Educational and Psychological Tests and Manuals. Washington, D.C.: American Psychological Association, cited in Martuza, V.R., 1977, p.151.
- Augstein, E.S. Harri-. 1976. An Investigation into Reading-for-Learning Problems: The Development of Two Awareness-Raising Tools. Centre for the Study of Human Learning, Brunel University, (typescript).
- Augstein, E.S. Harri-, and Thomas, L.F. 1976. The Kelly Repertory Grid as a Vehicle for Eliciting a Personal Taxonomy of Purposes in Reading. Centre for the Study of Human Learning, Brunel University, (typescript).
- Augstein, E.S. Harri-, and Thomas, L.F. 1984. "Conversational Investigations of Reading: the Self-Organized Learner and the Text." In Alderson, J.C., and Urquhart, A.H., (Eds.) 1984.
- Augstein, E.S. Harri-, Smith, M., and Thomas, L.F. 1982. Reading to Learn. Great Britain: Methuen.
- Barnitz, J.G. 1985. Reading Development of Nonnative Speakers of English. ERIC. Washington, D.C. (typescript).
- Bastien, S. 1983. "Investigación Básica y Elaboración de Materiales de Comprensión de Lectura." Estudios de Lingüística Aplicada, Número Especial, (enero de 1983). Mexico: CELE, UNAM, pp.80-103.
- Beaugrande, R. de. 1984. "Reading skills for foreign languages: a processing approach." In Pugh, A.K., and Ulijn, J.M. (Eds.). 1984.
- Berman, R.A. 1984. "Syntactic components of the foreign language reading process." In Alderson, J.C., and Urquhart, A.H. (Eds.). 1984.
- Biggs, J.B., and Collis, K.F. 1982. Evaluating the Quality of Learning: The SOLO Taxonomy (Structure of the Observed Learning Outcome). New York: Academic Press.
- Block, E. 1986. "The Comprehension Strategies of Second Language Readers." TESOL Quarterly. Vol. 20, No.3, pp.463-494.
- Bloom B.S. (Ed.) 1956. Taxonomy of Educational Objectives: Cognitive Domain. New York: McKay.
- Bray, J.H., and Maxwell, S.E. 1985. Multivariate Analysis of Variance. Sage University Paper series on Quantitative Applications in the Social Sciences, series no.07-054 Beverly Hills and London: Sage Publications.

- British Council. 1978. The Teaching of Comprehension, an Occasional Paper. ETIC Publications. London: The British Council.
- Brown, A.L., and Day, J.D. 1983. "Macrorules for Summarizing Texts: The Development of Expertise." Journal of Verbal Learning and Verbal Behaviour, Vol. 22, pp. 1-14.
- Brown, H.D., Yorio, C.A., and Crymes, R.H. (Eds.). 1977. Teaching and Learning ESL: Trends on Research and Practice. (On TESOL 1977). Washington, D.C.: TESOL.
- Bruder, M.N., and Henderson, R.T. 1986. Beginning Reading in English as a Second Language. ERIC. Washington, D.C. (typescript).
- Butler, C. 1985. Statistics in Linguistics. Oxford and New York: Basil Blackwell.
- Cao-Romero, L. 1983. El resumen como medio para evaluar la comprensión de lectura. Term project, M.A. in Applied Linguistics, CELE, UNAM, (typescript).
- Carrell, P.L. 1983. "Three Components of Background Knowledge in Reading Comprehension." Language Learning, Vol. 33, No. 2, pp. 183-207.
- Carrell, P.L. 1985. "Facilitating ESL Reading by Teaching Text Structure." TESOL Quarterly, Vol. 19, No.4, pp.727-752.
- Castañón, F. 1979. Reading Difficulties Encountered in Engineering. Talk given at MEXTESOL Conference, Cocoyoc, Mexico, (typescript).
- Cavalcanti, M.C. 1987. "Investigating FL Reading Performance through Pause Protocols." In Faerch, C., and Kasper, G. (Eds.). 1987.
- Chasan, M. 1981. "Comprension de Lectura: Una Comparación de Dos Metodologías." Estudios de Lingüística Aplicada, Vol.1, No.1, July 1981. Mexico: CELE, UNAM, pp.121-131.
- Chasan, M., and Williamson, M. 1979. "Field of Study-Learning Strategies; Are They Related?" Boston: TESOL, 1979, (typescript).
- Clarke, M.A. 1979. "Reading in Spanish and English: Evidence from Adult ESL Students." Language Learning, Vol.29, No.1, pp. 121-150.
- Clarke, M. 1980. "The Short Circuit Hypothesis of ESL Reading- or When Language Competence Interferes with Reading Performance." Modern Language Journal, Vol.64, No.2, pp.203-209.
- Coady, J. 1979. "A Psycholinguistic Model of the ESL Reader." In Mackay, R., Barkman, B. and Jordan, R.R. (Eds.). Reading in a Second Language. Rowley, Mass.: Newbury House.

- Cohen, A.D. 1984. "The use of mentalistic measures in determining LSP reading problems." In Pugh, A.K., and Ulijn, J.M. (Eds.). 1984.
- Cohen, A., and Hosenfeld, C. 1981. "Uses of Introspection/Retrospection in Learner-centered Research." Language Learning, Vol. 31, pp.285-313.
- Connaughton, B., and Flores, T. 1978. Teaching English for Academic Purposes in Mexico. Paper presented during the 1978 TESOL Annual Convention, Mexico City, 1978, (mimeo).
- Connor, U. 1984. "Recall of Text: Differences Between First and Second Language Readers." TESOL Quarterly, Vol. 18, No.2, pp. 239-256.
- Cooper, M. 1984. "Linguistic competence of practised and unpractised non-native readers in English." In Alderson, J.C., and Urquhart, A.H. (Eds.) 1984.
- Corder, S.P. 1981. "Strategies of Communication." Error Analysis and Interlanguage. Oxford: Oxford University Press.
- Coulthard, M. 1977. An Introduction to Discourse Analysis. London: Longman.
- Cowan, J.R. 1976. "Reading, Perceptual Strategies and Contrastive Analysis." Language Learning, Vol. 26, pp. 95-109.
- Cronbach, L.J. 1961. Essentials of Psychological Testing. New York: Harper and Row, cited in Davies, A. 1977 b., p.59.
- Cziko, G. 1980. "Language Competence and Reading Strategies: A Comparison of First and Second Language Oral Reading Errors." Language Learning, Vol.30, pp. 101-116.
- Dahlgren, L.O. 1984. "Outcomes of Learning." In Marton, F., Hounsell, D. and Entwistle, N. (Eds.). 1984.
- Davis, A. 1977 a. "Introduction." In Allen J.P.B., and Davies, A. (Eds.). 1977.
- Davies, A. 1977 b. "The Construction of Language Tests." In Allen J.P.B., and Davies, A. (Eds.) 1977.
- Davies, A. 1984. "Simple, Simplified and Simplification: What Is Authentic?" In Alderson, J.C., and Urquhart, A.H. (Eds.).1984.
- Devine, J. 1981. "Developmental Patterns in Native and Non-Native Reading Acquisition." In Hudelson, S. (Ed.) 1981.
- Entwistle, N.J. 1981. Styles of Learning and Teaching. Great Britain: John Wiley and Sons, Ltd. (Reprinted, 1983).
- Entwistle, N. 1985. "Ability Levels and Attainment Outcomes." in Pollitt, A., et.al. 1985.

- Entwistle, N., and Marton, F. 1984. "Changing Conceptions of Learning and Research." In Marton, F., Hounsell, D., and Entwistle, N. (Eds.). 1984.
- Entwistle, N.J., and Ramsden, P. 1983. Understanding Student Learning. London: Croom Helm.
- Faerch, C., and Kasper, G. (Eds.). 1987. Introspection in Second Language Research. Clevedon and Philadelphia: Multilingual Matters.
- Farr, R., Carey, R., and Tone, B. 1986. "Recent Theory and Research into the Reading Process: Implications for Reading Assessment." In Orasanu, J. (Ed.). 1986.
- Fawcett, R. 1979. "Reading Laboratories." In Lunzer, E., and Gardner, K. (Eds.). 1979.
- Flesch, R.F. 1948. "A new readability yardstick." Journal of Applied Psychology, Vol. 32, pp.221-233.
- Flesch, R.F. 1949. The Art of Readable Writing. New York: Harper and Row. (1974 Edition).
- Fransson, A. 1977. "On Qualitative Differences in Learning: IV-Effects of Intrinsic Motivation and Extrinsic Test Anxiety on Process and Outcome." Br.J.educ.Psychol., Vol. 47, pp.244-257.
- Fransson, A. 1984. "Cramming or Understanding? Effects of intrinsic and extrinsic motivation on approach to learning and test performance." In Alderson, J.C., and Urquhart, A.H. (Eds.). 1984.
- Gaito, J. 1970. "Repeated Measurements Designs and Counterbalancing." In Heermann, E.F., and Braskamp, L.A. (Eds.). 1970.
- Galicia, F. 1983. Estudio de las Necesidades de Lectura en Inglés de los Alumnos de la División de Estudios de Posgrado de la Facultad de Medicina de la UNAM. DLA, CELE, UNAM, Mexico, (typescript).
- Goodman, K.S. 1973 a. "Analysis or Oral Reading Miscues: Applied Psycholinguistics." In Smith, F. (Ed.). 1973.
- Goodman, K.S. 1973 b. "Psycholinguistic Universals of the Reading Process." In Smith, F. (Ed.). 1973.
- Grotjahn, R. 1987. "On the Methodological Basis of Introspective Methods." In Faerch, C., and Kasper, G. (Eds.). 1987.
- Grundin, H., et.al. 1981. "Cloze Procedure and Comprehension: an Exploratory Study across Three Languages." Journal of Research in Reading, Vol. 4, No.2, pp. 101-122.

- Haastrup, K. 1987. "Using Thinking Aloud and Retrospection to Uncover Learners' Lexical Inferencing Procedures." In Faerch, C., and Kasper, G. (Eds.). 1987.
- Hall, W.S., White, T.G., and Guthrie, L. 1986. "Skilled Reading and Language Development: Some Key Issues." In Orasanu, J. (Ed.). 1986.
- Harrison, C. 1979. "Assessing the Readability of School Texts." In Lunzer, E., and Gardner, K. (Eds.). 1979.
- Hatch, E., and Farhady, H. 1982. Research Design and Statistics for Applied Linguistics. Rowley, Massachusetts: Newbury House Publishers, Inc.
- Hauptman, P.C. 1981. "A Comparison of First and Second Language Reading Strategies." ITL Louvain, Vol. 51, pp. 37-57.
- Hays, W.L. 1973. Statistics for the Social Sciences. Great Britain: Holt, Rinehart and Winston. (2nd edition).
- Heermann, E.F., and Braskamp, L.A. (Eds.). 1970. Readings in Statistics for the Behavioral Sciences. New Jersey: Prentice-Hall, Inc.
- Henzell-Thomas, J. 1985. Learning from Informative Text: Prediction protocols as a means of accessing the interaction between top-down and bottom-up processes. Ph.D. Thesis: University of Lancaster, pp. 345-381.
- Hosenfeld, C. 1984. "Case studies of ninth grade readers." In Alderson, J.C., and Urquhart, A.H. (Eds.). 1984.
- Howatt, A.P.R. 1984. A History of English Language Teaching. Oxford: Oxford University Press. (2nd impression, 1985).
- Hudelson, S. (Ed.). 1981. Learning to Read in Different Languages. Washington, D.C.: Center for Applied Linguistics.
- Hutchinson, C. 1985. "What Makes Comprehension Questions Difficult in English and French?" In Pollitt, A., et.al. 1985.
- Ingram, E. 1975. English Language Battery Test. Edinburgh: Edinburgh University.
- Ingram, E. 1977. "Basic Concepts in Testing." In Allen, J.P.B., and Davies, A. (Eds.) 1977.
- Irwin, J.W. 1979. Text Structure and Comprehension: New Directions for Readability. Paper presented at the 1979 Annual Meeting of the National Reading Conference, San Antonio, Texas. ERIC. Washington, D.C. (typescript).
- Iversen, G.R., and Norpoth, H. 1976. Analysis of Variance. Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07-001, Beverly Hills and London: Sage Publications.
- Johnson, P. 1982. "Effects on Reading Comprehension of Building Background Knowledge." TESOL Quarterly, Vol.16, pp.503-516.

- Jones, E.E., and Gerard, H.B. 1967. Foundations of Social Psychology. New York: John Wiley and Sons.
- Kerlinger, F.N. 1967. Foundations of Behavioral Research. New York: Holt, Rinehart and Winston, Inc.
- Kintsch, W., and van Dijk, T.A. 1978. "Toward a Model of Text Comprehension and Production." Psychological Review, Vol. 85, No. 5, pp. 363-394.
- Kline, P. 1986. A Handbook of Test Construction. London: Methuen.
- Kozminsky, E., and Graetz, N. 1986. "First vs. second language comprehension: some evidence from text summarizing." Journal of Research in Reading, Vol.9, No.1, pp. 3-21.
- Laufer, B., and Sim, D.D. 1985. "Taking the Easy Way Out: Non-use and Misuse of Clues in EFL Reading." English Teaching Forum, April, 1985: pp. 7-20.
- Lautamatti, L. 1978. "Developing Materials for Teaching Reading Comprehension in a Foreign Language." In The Teaching of Comprehension, an Occasional Paper, ETIC Publications, London: The British Council, pp. 92-109.
- Lazarsfeld, P.F. 1949. "The American Soldier-an Expository Review." Publ.Opin.Quart., Vol.13, pp. 377-404, cited in Jones, E.E., and Gerard, H.B., 1967, pp. 34-35.
- Leong, C.K. 1984. "Coding Processes in Reading Alphabetic and Morphemic Scripts." In Pugh, A.K., and Ulijn, J.M., (Eds.).1984
- Lewis, D.G. 1967. Statistical Methods in Education. London: University of London Press Ltd. (3rd impression, 1972).
- Lindquist, E.F. 1940. Statistical Analysis in Educational Research. Boston: Houghton Mifflin.
- Lindquist, E.F. 1953. Design and Analysis of Experiments in Psychology and Education. Boston: Houghton Mifflin.
- Lunzer, E., and Gardner, K. (Eds.). 1979. The Effective Use of Reading. London: Heinemann.
- Lunzer, E., Waite, M., and Dolan, T. 1979. "Comprehension and Comprehension Tests." In Lunzer, E., and Gardner, K. (Eds.). 1979.
- Mackay, R., Barkman, B., and Jordan, R.R., (Eds.). 1979. Reading in a Second Language. Rowley, Mass.: Newbury House.
- Mackay, R., and Klassen, B. 1975. An Operational Framework for the Classification of Reading Comprehension Exercise Types. Research and Development Unit, Report No.3, mimeo, CELE UNAM, Mexico.
- Mackay, R., et.al. 1975 a. Basic Information Gathering Project. Research and Development Unit, Report No.1, mimeo, CELE, UNAM, Mexico.

- Mackay, R., et.al. 1975 b. A Project in English for Specific Purposes. Research and Development Unit, Report No. 2, mimeo, CELE, UNAM, Mexico.
- Mackay, R., et.al. 1975 c. A Trial Use of Cloze Test Procedure in the Assessment of Reading Comprehension. Research and Development Unit, Report No. 4, mimeo, CELE, UNAM, Mexico.
- Mackay, R., et.al. 1976 a. Desempeño del Alumno como Medio para Evaluar el Programa de Lectura de Comprensión en Inglés de los Estudiantes de la Licenciatura en la Facultad de Medicina Veterinaria y Zootecnia de la UNAM. Research and Development Unit, Report No. 6, mimeo, CELE, UNAM, Mexico.
- Mackay, R., et.al. 1976 b. A Study of the Predominant Reading Comprehension Difficulties as Demonstrated by Undergraduate Students at the UNAM. Research and Development Unit, Report No. 7, mimeo, CELE, UNAM, Mexico.
- Markham, P.L. 1987. "Rational Deletion Cloze Processing Strategies: ESL and Native English." System, Vol.15, No.3, pp. 303-311.
- Marton, F. 1976. "What does it take to learn? Some implications of an alternative view of learning." In Entwistle, N. (Ed.). 1976. Strategies for Research and Development in Higher Education. Amsterdam: Swets and Zeitlinger, cited in Biggs J.B., and Collis, K.F. 1982. p.3.
- Marton, F. 1981. "Phenomenography-Describing Conceptions of the World Around Us." Instructional Science, Vol.10, pp.177-200
- Marton, F., Hounsell, D., and Entwistle, N.J. (Eds.). 1984. The Experience of Learning. Great Britain: Scottish Academic Press.
- Marton, F., and Säljö, R. 1976 a. "On qualitative differences in learning: I - Outcome and Process." Br.J.educ.Psychol., Vol.46, pp. 4-11.
- Marton, F., and Säljö, R. 1976 b. "On qualitative differences in learning: II - Outcome as a function of the learner's conception of the task." Br.J.educ.Psychol., Vol.46, pp.115-127
- Martuza, V.R. 1977. Applying Norm-Referenced and Criterion-Referenced Measurement in Education. Boston: Allyn and Bacon.
- Meara, P. 1984. "Word recognition in foreign languages." In Pugh, A.K., and Ulijn, J.M. (Eds.). 1984.
- Moroney, M.J. 1951. Facts from Figures. Harmondsworth: Penguin. (Reprinted, 1954).
- Nash, T.W. 1983. Teaching Punctuation to Beginning ESL/EFL Students. Unpublished Master's Paper. Pittsburgh: University of Pittsburgh, Department of General Linguistics, cited in Bruder, M.N., and Henderson, R.T. 1986.

- Norusis, M.J. 1985. SPSS^X Advanced Statistics Guide. New York: McGraw-Hill.
- Nuttall, C. 1982. Teaching Reading Skills in a Foreign Language. London: Heinemann Educational Books, Ltd. (Reprinted, 1983).
- Oller, J.W. Jr. 1979. Language Tests at School. London: Longman.
- Orasanu, J. (Ed.). 1986. Reading Comprehension: From Research to Practice. New Jersey and London: Lawrence Erlbaum Associates, Publishers.
- Pearson, P.D. (Ed.). 1984. Handbook of Reading Research. New York and London: Longman.
- Peel, E.A. 1971. The Nature of Adolescent Judgement. London: Staples Press, cited in Entwistle, N. 1985. P. 6.
- Perkins, K., and Jones, B. 1985. "Measuring Passage Contribution in ESL Reading Comprehension." TESOL Quarterly, Vol. 19, No. 1, pp. 137-153.
- Perry, W.G. 1983. "Forward." In Entwistle, N.J., and Ramsden, P. 1983.
- Pollitt, A., et.al. 1985. What Makes Exam Questions Difficult? (Research Reports for Teachers, No.2). Edinburgh: Scottish Academic Press.
- Polunin, N. (Ed.). 1980. Growth Without Ecodisasters? London and Basingstoke: The Macmillan Press Ltd.
- Pugh, A.K. 1978. Silent Reading. London: Heinemann Educational Books.
- Pugh, A.K. and Ulijn, J.M. (Eds.). 1984. Reading for Professional Purposes. London: Heinemann Educational Books.
- Reichmann, W.J. 1964. Use and Abuse of Statistics. Harmondsworth: Penguin.
- Richards, J., Platt, J., and Weber, H. 1985. Longman Dictionary of Applied Linguistics. Great Britain: Longman.
- Rigg, P. 1977. "The Miscue ESL Project." In Brown, H.D., Yorio, C.A., and Crymes, R.H. (Eds.). 1977.
- Robson, C. 1973. Experiment Design and Statistics in Psychology. Harmondsworth: Penguin.
- Rowntree, D. 1981. Statistics Without Tears. Harmondsworth: Penguin. (Reprinted, 1985).
- Ryan, P. 1987. El Resumen y el Proceso de Redacción. Reporte de Investigación DC/I-03, DLA, CELE, UNAM, Mexico, (typescript).

- Samuels, S.J., and Kamil, M.L. 1984. "Models of the Reading Process." In Pearson, P.D. (Ed.). 1984.
- Savignon, S.J. 1983. Communicative Competence: Theory and Classroom Practice. Massachusetts: Addison-Wesley Publishing Company.
- Schroder, H.M., Driver, M.J., and Streufert, S. 1967. Human Information Processing. New York: Holt, Rinehart, and Winston, cited in Biggs, J.B., and Collis, K.F. 1982. p.13.
- Singer, M. 1982. "Comparing Memory for Natural and Laboratory Reading." Journal of Experimental Psychology: General. Vol. 111, No. 3, pp. 331-347.
- Smith, F. (Ed.). 1973. Psycholinguistics and Reading. New York: Holt, Rinehart and Winston.
- Smith, F. 1978. Reading. Cambridge: Cambridge University Press.
- Smith, H.W. 1975. Strategies of Social Research. The Methodological Imagination. Englewood Cliffs, New Jersey, Prentice-Hall, Inc.
- Spector, P.C. 1981. Research Designs. Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07-023, Beverly Hills and London: Sage Publications.
- Statman, S. 1987. "Obstacles to Access: an Investigation into the Perceptual Strategies of the Non-native Learner of English." System. Vol. 15, No. 3, pp. 289-301.
- Steffensen, M.S., and Joag-Dev, C. 1984. "Cultural knowledge and reading." in Alderson, J.C., and Urquhart A.H. (Eds.). 1984.
- Stern, H.H. 1983. Fundamental Concepts of Language Teaching. Great Britain: Oxford University Press. (3rd impression, 1984).
- Strong, M.F. 1980. "The International Community and the Environment." In Polunin, N. (Ed.). 1984.
- Svensson, L. 1977. "On qualitative differences in learning: III - study skill and learning." Br.J.educ.Psychol., Vol.47, pp. 233-243.
- Svensson, L. 1984. "Skill in Learning." In Marton, F., Hounsell, D., and Entwistle, N. (Eds.). 1984.
- Taylor, W.L. 1953. "'Cloze Procedure': A New Tool for Measuring Readability." Journalism Quarterly. Vol.30, pp. 415-433.
- Thomas, L.F., and Augstein, E.S. Harri-. 1972. "An Experimental Approach to the Study of Reading as a Learning Skill." Research in Education, Vol. 8, pp. 28-46.

- Thomas, L.F., and Augstein, E.S. Harri-. 1976 a. Conversational Investigations of Reading. New Insights Into Cognition, Centre for the Study of Human Learning, Brunel University, (typescript).
- Thomas, L.F., and Augstein, E.S. Harri-. 1976 b. The Self-organised Learner and the Printed Word. Report to the Social Science Research Council. Centre for the Study of Human Learning, Brunel University, (typescript).
- Thomas, L.F., and Augstein, E.S. Harri-. 1976 c. "The Teaching of Reading-for-Learning." The Self-organised Learner and the Printed Word, Paper 1. Centre for the Study of Human Learning, Brunel University, (typescript).
- Thomas, L.F., Augstein, E.S. Harri-, and Farnes, N.C. 1973. "Reading-for-Learning: The Anatomy of a Research Project." In The Road to Effective Reading. Proceedings of the 10th Annual Conference of UKRA. London: Ward Lock Educational.
- Tudor, I. 1986. "Advance Organisers as Adjuncts to L2 Reading Comprehension." Journal of Research in Reading, Vol.9, No. 2, Sept. 1986, pp. 103-115.
- Ulijn, J.M. 1984. "Reading for professional purposes: psycholinguistic evidence in a crosslinguistic perspective." in Pugh, A.K., and Ulijn, J.M. (Eds.). 1984.
- Ulijn, J.M. 1985. "A present state of LSP reading research as reflected by a recent symposium." In Ulijn, J.M., and Pugh, A.K. (Eds.). 1985.
- Ulijn, J.M., and Kempen, G.A.M. 1976. "The Role of the First Language in Second Language Reading Comprehension. Some Experimental Evidence." In Nickel, G. (Ed.). Proceedings of the 4th International Congress of Applied Linguistics, Vol, 1, Hochschulverlag, Stuttgart, pp. 495-507, cited in Alderson, J.C. 1984. p. 12.
- Ulijn, J.M. and Pugh, A.K. (Eds.). 1985. Reading for Professional Purposes: Methods and Materials in Teaching Languages. Leuven: ACCO.
- Urquhart, A.H. 1978. "Comprehension: The Discourse Analysis View: Meaning in Discourse." In The Teaching of Comprehension, an Occasional Paper, ETIC Publications, London, The British Council.
- Urquhart, A.H. 1984. "The effect of rhetorical ordering on readability." In Alderson, J.C., and Urquhart, A.H. (Eds.). 1984.
- Winograd, P.N. 1984. "Strategic Difficulties in Summarizing Texts." Reading Research Quarterly. Vol. 19, No.4, pp.404-425.

- Yorio, C.A. 1971. "Some Sources of Reading Problems for Foreign Language Learners." Language Learning, Vol.21, pp.107-115.
- Zuck, L.V., and Zuck, J.G. 1984. "The main idea: specialist and non-specialist judgements." In Pugh, A.K., and Ulijn, J.M. (Eds.). 1984.